



Combination possibility and deep learning model as clinical decision-aided approach for prostate cancer

Health Informatics Journal

1–18

© The Author(s) 2019

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1460458219855884

journals.sagepub.com/home/jhi**Okyaz Eminaga** 

Stanford University School of Medicine, USA; University Hospital of Cologne, Germany

Omran Al-Hamad

University Hospital of Cologne, Germany

Martin Boegemann

University Hospital Muenster, Germany

Bernhard Breil* 

Niederrhein University of Applied Science, Germany

Axel Semjonow*

University Hospital Muenster, Germany

Abstract

This study aims to introduce as proof of concept a combination model for classification of prostate cancer using deep learning approaches. We utilized patients with prostate cancer who underwent surgical treatment representing the various conditions of disease progression. All possible combinations of significant variables from logistic regression and correlation analyses were determined from study data sets. The combination possibility and deep learning model was developed to predict these combinations that represented clinically meaningful patient's subgroups. The observed relative frequencies of different tumor stages and Gleason score GIs changes from biopsy to prostatectomy were available for each group. Deep learning models and seven machine learning approaches were compared for the classification performance of Gleason score changes and pT2 stage. Deep models achieved the highest F1 scores by pT2 tumors (0.849) and GIs change (0.574). Combination possibility and deep learning model is a useful decision-aided tool for prostate cancer and to group patients with prostate cancer into clinically meaningful groups.

Keywords

classification model, combination and learn model, deep learning, pathology, prediction model, prostate cancer, risk classification, SEER, wide learning

*Contributed equally as senior authors.

Corresponding author:

Okyaz Eminaga, Department of Urology, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305-5118, USA.

Email: ok yaz.eminaga@gmail.com

Introduction

Prostate cancer (PCa) is one of the most diagnosed malignancies in men and the third reason for cancer-associated death in the United States.¹ Most patients with clinically organ-confined PCa are treated with radical prostatectomy (RPE) or radiotherapy, which provides excellent cancer control.² However, there is no consensus regarding the optimal management of locally advanced PCa.² The prediction of the pathologic stage of PCa before an intervention enables improved patient counseling and clinical decision making for treatment planning and risk stratification for novel clinical trials for those patients with more advanced and aggressive PCa. Recent studies have published algorithms and nomograms predicting the pathologic stage of patients with clinically localized PCa or Gleason score upgrading.^{3–10} However, these prediction models are not frequently used due to limitations in usability and applied computational approaches. The recent advance in artificial intelligence (AI) and computational capabilities facilitates nowadays robust pattern recognitions and data structure determination in large data sets, imaging, and genetics. Although the application of AI in medicine remains in its early stages, the integration of AI in medicine is opening a new avenue in disease care management. In the past, some studies have introduced different prediction models for advanced PCa using conventional machine learning.^{11–13} The multi-layer neural networks approach, also called as a deep neural network, is one of the deep learning approaches, and it has demonstrated very accurate results in recognizing images and determining genetic variations. Whether the deep neural network approach is applicable in developing prediction models for PCa is not clear. This study will stress this question and introduce a proof of concept for utilizing deep learning approaches to predict pathologic outcomes. Here, we are going to develop a prediction model for pathologic outcomes using preoperative data, unsupervised clustering, followed by supervised machine learning, and deep learning approaches. We aim to support the clinical decision by providing decision-aiding tools which can learn how to predict the outcome on changing and different data sets with high accuracy.

Material and methods

The development procedure of the classification model for PCa consists briefly of four steps (Figure 1). First, we utilized four different data sets representing the different conditions of data acquisitions in epidemiologic studies: prospective data acquisition and storage in an institutional data set, retrospective data acquisition and storage in an institutional data set, the single source information system,^{14,15} and a national epidemiologic cancer registry. In this way, we limited the risk of selection bias that can occur in each of these data sets. We then extracted data related to clinicopathological data before and after the removal of the prostate as a consequence of PCa for each patient. The second step performs multinomial multivariate regression analyses and Spearman correlation on variables to determine the associations between preoperative parameters and tumor progression. The third step determines the number of all possible combinations of these variables *in silico* and identifies these combinations in our data set. The data sets will then be unsupervised clustered by the number of combinations seen on the real data sets. The final step aims at the development of optimal deep neural network models to predict the cluster “the new class” by an arbitrary definition of the dimensions of hidden layers of neural networks. Furthermore, we assumed that the aggressiveness of the tumor is independent of age at diagnosis and race. Therefore, these features (i.e. age and race) were defined as noisy features for the deep neural network. Both these features were not considered for the combination and clustering analyses.

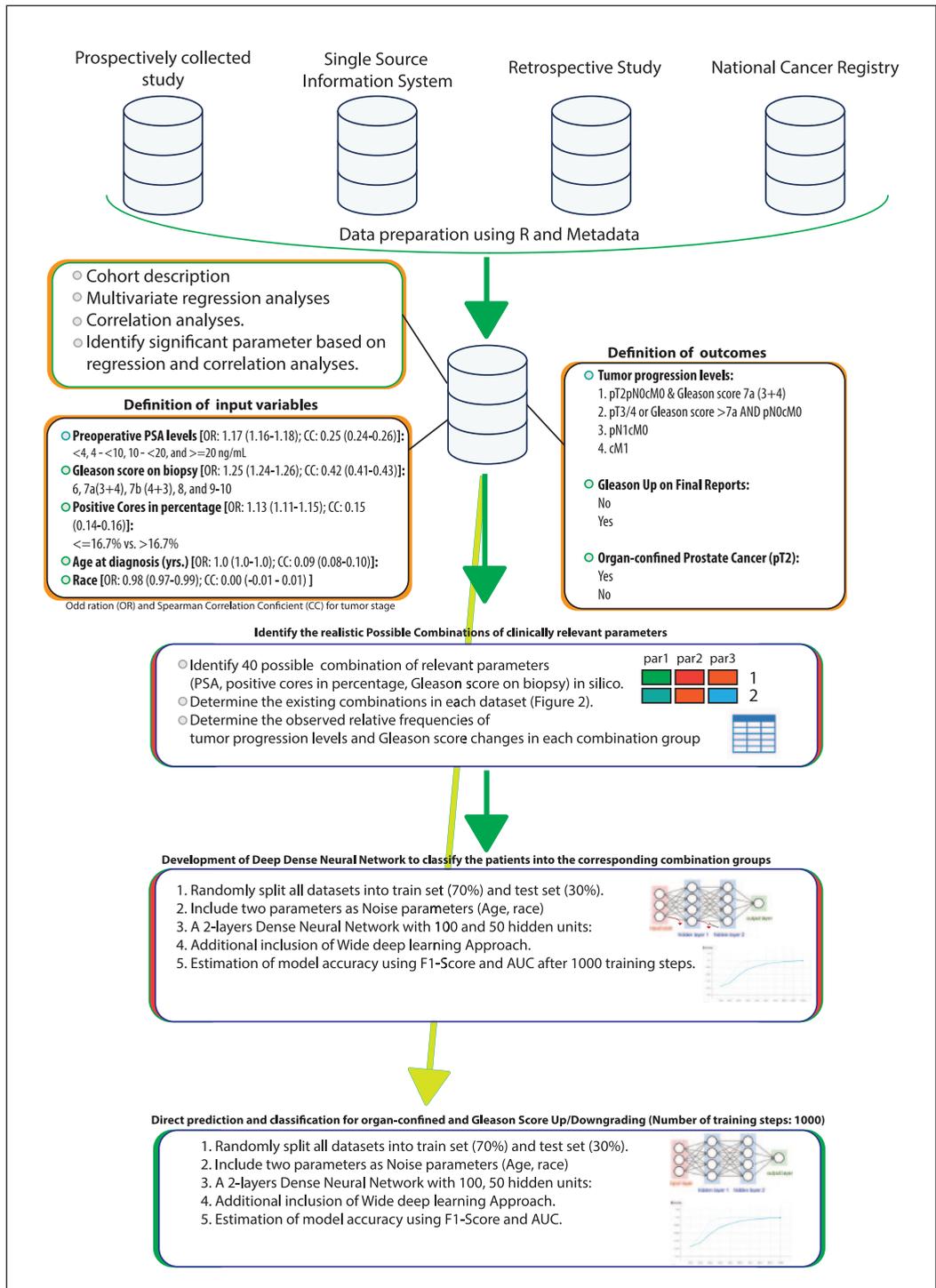


Figure 1. The workflow of the combination and deep model for prostate cancer.

Table 1. The variable definition.

Variables	Categorization
Race	White, Black, others
Age at diagnosis, (years)	63> vs ≤63
Preoperative serum PSA level (ng/mL)	<4, 4–10, 10.1–20, >20
Pathological node stage	pN0 vs pN1
Clinical metastasis stage	cM0 vs cM1
Pathological Gleason score on biopsy and on prostatectomy specimens	6 vs 7a (3 + 4) vs 7b (4 + 3) vs 8 vs 9–10
Number of positive cores at biopsy in percentage	16.7%> vs ≤16.7%

PSA: prostate-specific antigen.

Variable definition

Table 1 lists the variables considered for this study. Patient information included race (i.e. White, Black, others), age at diagnosis (years), preoperative serum prostate-specific antigen (PSA) level (ng/mL) as continuous and categorized variables (<4, 4–10, 10.1–20, >20 ng/mL), pathological tumor stage (pT2, pT3a, pT3b, pT4), pathological node stage (pN0, pN1), clinical metastasis stage (cM0, cM1), pathological Gleason score¹⁶ on biopsy and on prostatectomy specimens (6 vs 7a (3 + 4) vs 7b (4 + 3) vs 8 vs 9–10), number of positive cores at biopsy, and the total number of biopsy cores. The Gleason score was divided according to the recommendation of International Society of Urological Pathology. The parameter “Age at diagnosis” was dichotomized using the median split. We calculated the ratio of positive cores to the total number of biopsy cores taken by the prostate biopsy. The ratio of positive cores was categorized by considering the 25th percentile as a cutoff. Patients were divided into four groups as outcome groups by PSA, the TNM stage, and the pathologic final Gleason score status: patients with pT2pN0cM0 PCa and PSA levels < 10 ng/mL and final Gleason score ≤ 7a; patients with final Gleason score > 7a or pT3a/b or pT4 or PSA ≥ 10 ng/mL and no evidence of metastasis; patients with regional lymph node metastasis; and patients with distant metastases. The data set was pseudonymized during the whole processing steps.

Data extraction

Data acquisition was performed according to the precepts of the Helsinki Declaration and German data privacy regulation.

Prospectively collected data. We utilized clinicopathological prospectively collected data of 533 patients from a recent study, evaluating the variation of the tumor location between men diagnosed at initial biopsy and those diagnosed at repeat biopsy.¹⁷ The required data were extracted from an Excel spreadsheet and stored in comma-separated file format (CSV). The data processing is then performed using the R Console (R Foundation for Statistical Computing, Vienna, Austria) and a general metadata which facilitates data merging with the following data set.

Extraction from the electronic health record. We applied the biopsy report generator introduced by Breil et al.¹⁸ and the final pathology report developed by Eminaga et al.¹⁹ to directly extract the clinicopathological data from Hospital Information System (HIS). The required data were then stored in CSV file format and processed using a proper metadata and the R Console. Data from

biopsy and pathology reports were merged by patient identity number (PID), which was removed from the merged data.

Retrospective data. Retrospectively collected data of 455 patients who underwent total removal of the prostate in the University Hospital of Cologne in the period between 2004 and 2008 were considered. The data were stored by database storage software (Microsoft Access) and extracted by Structured Query Language (SQL) queries. The extracted data were stored in CSV file format and processed in R Console.

National registry data. The Surveillance Epidemiology and End Results (SEER) 18 registries were used for this study. SEER consists of population-based cancer registries, representing approximately 28 percent of the US population, and provides data related to basic demographics, tumor site, histology, stage, grade, and treatments performed. The study included only men ≥ 35 years of age diagnosed between 2004 and 2014 with histologically confirmed adenocarcinoma (International Classification of Diseases for Oncology, third edition, code 8140) of the prostate (site code 61.9). All cases ($n=44,321$) were staged based on the American Joint Committee on Cancer (AJCC) Cancer Staging Manual, sixth edition, with either radiographic or pathologic confirmation of metastatic disease.²⁰ Cases diagnosed by autopsy or death certificate only were excluded. Following SEER coding guidelines, data regarding age, race, Gleason score on biopsy and prostatectomy specimens, and pathologic AJCC-based T, N stages, and the clinical M stage were acquired at the time of diagnosis. Information about the pathological AJCC T stage was derived from RPE. Information about the pathological AJCC N stage was derived from any microscopic assessment of regional nodes with RPE/prostate biopsy or from autopsies in men who had been diagnosed with PCa before death (CS lymph nodes eval codes 2 or 3). The PSA measurement corresponded to the highest PSA value recorded before diagnostic prostate biopsy or treatment. The population was divided by race into White, African American, and others or unknown. Cases with total removal of the prostate (surgery site codes 50 or 70) were identified. Other forms of local therapy, including incomplete removal of the prostate (surgery site code 30), focal therapy (FT; e.g. cryotherapy, laser, hyperthermia; surgery site codes 10–17 or 24–26), or transurethral resection of the prostate (TURP; surgery site codes 19–26) were excluded. Patients with unknown therapy were also excluded. The extracted data were transferred and processed using the proper metadata using R Console.

Data analyses

After data preparation, we performed multinomial multivariate regression analyses and Spearman correlation coefficients to evaluate the odds ratios of selected variables (i.e. PSA; GIs at biopsy, race, and age at diagnosis; ratios of positive/negative cores) for different tumor stages and their correlation to tumor progression. The q value (false discovery rate (FDR)-adjusted p value) was estimated for comparative analyses. All statistical tests were two-sided, and the level of statistical significance was set at $q \leq 0.05$.

We calculated the number of all possible combinations of significant and categorized variables using the following equation

$$c = \prod_{k=1}^n |a_k|$$

where c is the number of combination possibilities, k is the index of variable a , n denotes the total number of variables, and $|a_k|$ is the number of features for each variable. After that, we determined

the existing (real) combinations in our data set by applying the duplication removal algorithm to identify realistic combination possibilities. After that, the data set was unsupervised clustered using the hierarchical clustering on the significant categorized variables. We repeated the duplication removal algorithm to verify that the data set is correctly clustered into the realistic combination possibilities. All analyses were performed with the R statistical package system (R Foundation for Statistical Computing).

Development of prediction models

We trained three models to predict the correct cluster for each patient in our data set: a wide model, which is a linear model with a wide set of sparse and crossed feature column as already described by Cheng et al.;²¹ a deep feed-forward neural network; and a model combining the wide and deep neural network. For the wide model, the crossed feature columns were between categorized PSA levels and the categorized parameter for ratios of positive cores and between categorized PSA levels and the Gleason score on biopsy. In this way, wide models with crossed feature columns enable memorizing sparse interactions between features effectively.²¹ For the deep model, we arbitrary defined different hidden units for deep neural networks; each hidden unit has two neural layers. The dropout regularization technique with a dropout rate of 0.2 was utilized to reduce the risk of overfitting²² by randomly selecting nodes to be dropped out with a given probability (in our study: 20%) of each weight update cycle. A centered bias variable is estimated for each cluster. The optimization algorithm “Adaptive Moment Estimation” (Adam) was used to compute adaptive learning rates for each parameter, thereby optimizing the neural network models. We preferred Adam due to its popularity in the field of deep learning and because Adam achieved better results in short training period compared to other approaches.²³ For the model optimization, we defined an initial learning rate of 0.001, a beta1 value of 0.9, beta2 value of 0.999, epsilon of value 1e-09. Rectified linear unit (ReLU) activation function was used to regulate the firing rate of neurons in the layer. For each analysis, a training set (70%) and a test set (30%) were randomly generated from the study data set by considering that the overall distribution of endpoints has met between train and test sets (Supplemental Table 1). All models were trained on the training set and evaluated on the test set. The training set was shuffled by each training step, and the batch size was defined as the number of training cases. The training steps were limited to 1000 steps to avoid the overfitting risks of these models. Through the evaluation (validation) process, prediction and classification accuracies, as well as precision, were quantified with the area under the curve (AUC), classification accuracy, precision, recall and F-measure (F1 score). Input data have all significant variables identified by the data analyses’ section. Furthermore, age at diagnosis and race were added as noise parameters into our input data to reduce the overfitting risks of our models, since the predictive value of these parameters for advanced PCa and Gls upgrading is controversial.

Moreover, we evaluated the analyses between deep neural networks and wide-deep neural networks model, random forest analyses, adapted boost, naïve Bayes and k-nearest neighbor’s algorithm, multivariate logistic regression analyses, classification tree and supported vector machine. Here, a training set was generated from SEER data sets; a test set was generated from SEER data sets. For comparison analyses, we classified the study cohort by presence of organ-confined PCa (pT2) or Gls upgrading status.

We utilized Python 2.7 (Python Software Foundation, Wilmington, USA), and Tensorflow (Google Inc., Mountain View, USA) for developing the models. All analyses were performed on a processor Intel i7 with RAM 32 GB and GPU NVIDIA™ GeForce GTX 1080 Ti with 11 GB VRAM.

Results

Table 2 shows an overview of the relevant clinicopathological information from each of the four different data sets. Overall, the median age at diagnosis was 63 years (interquartile range (IQR): 57–67 years). In total, 65 percent of men who underwent RPE had PSA levels between 4 and 10 ng/mL. The median biopsy cores were 12 (IQR: 12–12), and the 25th percentile of positive cores was 2. The 25th percentile of the positive cores ratio was 16.7 percent. A total of 70 percent of cases had positive cores in more than 16.7 percent of total biopsy cores. After surgery, 59.7 percent of men had locally advanced PCa. Loco-regional lymph node metastases were observed in 6.2 percent of cases. Only 99 (0.04%) men who underwent RPE had distant metastases. In the multivariate multinomial regression analyses and correlation analyses, categorized PSA levels, GIs by prostate biopsy, and categorized positive/negative cores given in percentage were identified as significant parameters. However, age at diagnosis and race were not the significant predictors in multivariate multinomial regression and correlation analyses. In silico, we identified 40 combination possibilities of these significant parameters. In Muenster's data set, we determined 38 possibilities for combining the significant parameters, whereas the Cologne's data set had 30 combination possibilities. The SEER database included data covering all combination possibilities. Figure 2 shows the Venn diagrams for intersections between these data sets. Table 3 shows the observed relative frequencies of different tumor progression levels and the Gleason upgrading for each possible combination (cluster). We found that certain clusters are remarkably associated with increased risk for advanced tumors or GIs scores upgrading in the prostatectomy pathology report.

The deep neural models recognized the corresponding cluster for each case with 100 percent accuracy, when age at diagnosis, race, categorized preoperative PSA levels and GIs score, and the percent of positive cores as Boolean parameter were considered as input data. Figure 3 exhibits the training progression for each model. The training duration varied between different deep neural models. The wide and deep neural network model achieved similar results, but with prolonged training periods.

Table 4 shows the results of the classification performance of different machine learning approaches. The deep models achieved higher accuracy than other approaches. By comparison analyses, the deep models achieved higher accuracies than other approaches. The deep models achieved the best F1 scores to classify patients by presence of pT2 tumors (0.850) and GIs changes (0.574) followed by the logistic regression (0.752 for pT2 tumors; 0.532 for GIs changes) and naïve Bayes for pT2 tumors (0.748) and random forest classification for GIs changes (0.530).

Discussion

The current model can successfully identify subsets of patients with high risk for advanced PCa or risk for GIs upgrading with high accuracy. Moreover, the current model can be fed with data from different data sources (retrospective, prospectively collected, single source information system, and cancer registry), representing real situations of data mining in clinical research. The introduced model is feasible to manage and reuse these data after applying metadata. Deep learning approach has received further attention in recent years after its successful application in image and object recognition and has been used nowadays for different research and commercial purposes.²⁴ However, our work bridges the gap of utilizing the deep learning approach in classifying patients with cancer according to their preoperative features to determine the observed relative frequency of outcomes. We preferred categorizing our input parameters to identify all possible combination of these parameters with each other. In clinical routine, the categorization of clinical data for risk estimation has been proven as a successful and decision-friendly tactic to help physicians in

Table 2. The cohort characteristics.

Characteristics	All	Retrospective	Prospectively collected	Single source information system	SEER National Cancer Registry
Population, n (%)	44,321 (100%)	455 (1.9)	511 (2.1)	133 (0.5)	43,341 (95.5)
Age at diagnosis, years, median (IQR)	62 (56–66)	65 (60–69)	64 (60–68)	66 (60–71)	62 (56–66)
Positive cores, median (IQR)	4 (2–6)	3 (2–5)	2 (1–3)	4 (2–5)	5 (3–7)
Total cores, median (IQR)	12 (12–12)	8 (6–12)	8 (6–8)	10 (10–12)	12 (12–13)
Positive cores in percentage					
≤16.7%	12,002 (27.1)	123 (27.0)	177 (34.6)	31 (23.3)	11,671 (27.0)
>16.7%	32,319 (72.9)	332 (73.0)	334 (65.4)	102 (76.7)	31,551 (73.0)
Race					
White, n (%)	35,931 (81.1)	455 (100)	511 (100)	133 (100)	34,832 (80.6)
Black, n (%)	5750 (13.0)	0 (0)	0 (0)	0 (0)	5750 (13.3)
Others, n (%)	2275 (5.1)	0 (0)	0 (0)	0 (0)	2275 (5.3)
Unknown, n (%)	365 (0.8)	0 (0)	0 (0)	0 (0)	365 (0.8)
PSA levels, ng/mL, n (%)					
<4	5950 (13.5)	27 (5.9)	36 (7.0)	7 (5.3)	5880 (13.6)
4–<10	29,749 (67.1)	258 (56.7)	299 (58.5)	71 (53.4)	29,121 (67.4)
10–<20	6218 (14.0)	111 (24.4)	133 (26.0)	35 (26.3)	5939 (13.7)
≥20	2404 (5.4)	59 (13.0)	43 (8.4)	20 (15.0)	2282 (5.3)
Gleason score on biopsy, n (%)					
6	17,189 (3.9)	255 (56.0)	249 (48.7)	67 (50.4)	16,618 (38.5)
7 (3 + 4)	14,392 (32.5)	70 (15.4)	143 (28.0)	39 (29.3)	14,140 (32.7)
7 (4 + 3)	6165 (13.9)	64 (14.1)	66 (12.9)	17 (12.8)	6018 (13.9)
8	4206 (9.5)	46 (10.1)	26 (5.1)	8 (6.0)	4126 (9.5)
9–10	2369 (5.3)	20 (4.4)	27 (5.3)	2 (1.5)	2320 (5.4)
Final Gleason score, n (%)					
6	11,095 (25.0)	248 (54.5)	113 (22.1)	21 (15.8)	10,713 (24.8)
7 (3 + 4)	20,078 (45.3)	88 (19.3)	197 (38.6)	67 (50.4)	19,726 (45.6)
7 (4 + 3)	7786 (17.6)	40 (8.8)	131 (25.6)	27 (20.3)	7588 (17.6)
8	2331 (5.3)	59 (13.0)	15 (2.9)	4 (3.0)	2253 (5.2)
9–10	3031 (6.8)	20 (4.4)	55 (10.8)	14 (10.5)	2942 (6.8)
Tumor stage, n (%)					
pT2	31,816 (71.8)	314 (69.0)	271 (53.1)	83 (62.4)	31,148 (72.1)
pT3a	7428 (16.8)	49 (10.8)	153 (29.9)	30 (22.6)	7196 (16.6)
pT3b	3756 (8.4)	75 (16.5)	60 (11.7)	20 (15.0)	3601 (8.3)
pT4	1321 (3.0)	17 (3.7)	27 (5.3)	0 (0)	1277 (3.0)
Lymph node status, n (%)					
N0	42,538 (96.0)	399 (87.7)	481 (94.1)	122 (91.7)	41,536 (96.1)
N1	1728 (3.9)	56 (12.3)	30 (5.9)	11 (8.3)	1631 (3.8)
NX	55 (0.1)	0 (0)	0 (0)	0 (0)	55 (0.1)
Metastasis, n (%)					
M0	44,212 (99.8)	444 (97.6)	511 (100)	133 (100)	43,124 (99.8)
M1a/b/c	109 (0.2)	11 (2.4)	0	0	98 (0.2)

Table 2. (Continued)

Characteristics	All	Retrospective	Prospectively collected	Single source information system	SEER National Cancer Registry
Tumor progression levels, n (%)					
1. pT2 and Gleason score 7a (3 + 4)	25,677 (57.9)	260 (57.1)	213 (41.7)	65 (48.9)	25,139 (58.1)
2. pT3/4 or Gleason score \geq 7b and pN0cM0	16,864 (38.0)	134 (29.5)	268 (52.4)	57 (42.9)	16,405 (38.0)
3. pNI	1671 (3.8)	50 (11.0)	30 (5.9)	11 (8.2)	1580 (3.7)
4. cMI	109 (0.3)	11 (2.4)	0 (0)	0 (0)	98 (0.2)
Gleason up/downgrading, n (%)					
Yes	20,125 (45.4)	210 (46.2)	292 (57.1)	70 (52.6)	19,553 (45.2)
No	24,196 (54.6)	245 (53.8)	219 (42.9)	63 (47.4)	23,669 (54.8)

SEER: Surveillance Epidemiology and End Results; PSA: prostate-specific antigen; IQR: interquartile range.

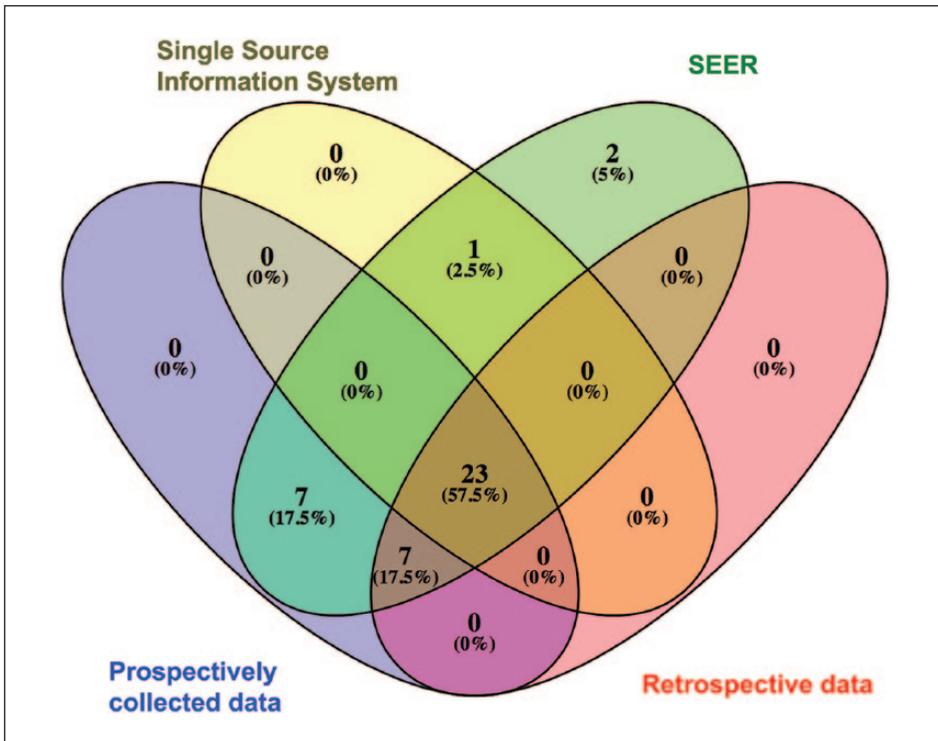


Figure 2. Venn diagram representing the number and portions of possible combinations between four data sets.

For instance, SEER data set has 40 combinations (color: green), whereas the prospectively collected data includes 37 combinations (color: lilac). By focusing on the intersection between these data sets, we found that the SEER data set has three combinations that are not available by Lila data set. Moreover, 23 (55%) possible combinations are available in all data sets.

Table 3. The observed relative frequencies of tumor progression levels and Gleason score changes for each patient group defined by PSA, positive cores in percentage, and Gleason score on biopsy.

ID	Combination Input parameters			Tumor progression					Gleason change		
	Gleason score on biopsy	Positive cores in percentage	PSA level (ng/mL)	n (%)	Level 1	Level 2	Level 3	Level 4	No	Upgrade	Downgrade
0	7a	>16.7	4-<10	8098 (18.3)	62.41%	35.96%	1.49%	0.14%	70.61%	18.62%	10.77%
1	6	>16.7	≥20	261 (0.6)	54.02%	40.61%	4.98%	0.38%	28.35%	71.65%	0.00%
2	6	>16.7	4-<10	7170 (16.2)	80.32%	19.25%	0.35%	0.08%	45.93%	54.07%	0.00%
3	6	≤16.7	≥20	181 (0.4)	59.67%	37.57%	2.21%	0.55%	44.20%	55.80%	0.00%
4	6	>16.7	<4	1619 (3.7)	87.40%	12.35%	0.25%	0.00%	56.95%	43.05%	0.00%
5	9-10	>16.7	≥20	333 (0.8)	2.10%	60.36%	35.74%	1.80%	73.87%	0.00%	26.13%
6	6	≤16.7	4-<10	4755 (10.7)	87.59%	12.30%	0.08%	0.02%	60.27%	39.73%	0.00%
7	6	>16.7	10-<20	960 (2.2)	62.50%	35.63%	1.77%	0.10%	32.19%	67.81%	0.00%
8	8	>16.7	10-<20	737 (1.7)	9.77%	73.41%	16.15%	0.68%	25.51%	22.80%	51.70%
9	7b	>16.7	10-<20	951 (2.1)	17.88%	71.19%	10.52%	0.42%	49.32%	19.98%	30.70%
10	6	≤16.7	10-<20	796 (1.8)	78.14%	20.73%	1.01%	0.13%	53.89%	46.11%	0.00%
11	7a	>16.7	<4	1350 (3.0)	70.30%	28.67%	0.96%	0.07%	70.89%	14.89%	14.22%
12	7a	>16.7	≥20	561 (1.3)	27.99%	57.75%	13.19%	1.07%	58.82%	38.32%	2.85%
13	7a	>16.7	10-<20	1556 (3.5)	43.57%	51.41%	4.82%	0.19%	64.01%	28.34%	7.65%
14	9-10	≤16.7	≥20	46(0.1)	10.87%	71.74%	17.39%	0.00%	17.39%	36.96%	45.65%
15	8	>16.7	≥20	455 (1.0)	10.55%	71.87%	17.14%	0.44%	53.19%	22.20%	24.62%
16	8	>16.7	4-<10	3245 (7.3)	31.25%	63.94%	4.62%	0.18%	45.33%	11.83%	42.84%
17	8	≤16.7	4-<10	741 (1.7)	40.89%	57.62%	1.35%	0.13%	43.72%	8.64%	47.64%
18	8	≤16.7	<4	113 (0.3)	53.98%	46.02%	0.00%	0.00%	34.51%	7.96%	57.52%
19	7b	≤16.7	4-<10	1945 (4.4)	73.42%	25.86%	0.72%	0.00%	66.43%	16.20%	17.38%
20	9-10	>16.7	4-<10	1132 (2.6)	7.86%	73.23%	17.76%	1.15%	59.10%	0.00%	40.90%

Table 3. (Continued)

ID	Combination Input parameters			Tumor progression					Gleason change		
	Gleason score on biopsy	Positive cores in percentage	PSA level (ng/mL)	n (%)	Level 1	Level 2	Level 3	Level 4	No	Upgrade	Downgrade
21	7a	≤16.7	<4	436 (1.0)	80.05%	18.81%	1.15%	0.00%	62.16%	13.07%	24.77%
22	7a	≤16.7	10-<20	370 (0.8)	59.73%	37.84%	2.43%	0.00%	62.97%	24.32%	12.70%
23	7b	≤16.7	10-<20	162 (0.4)	24.69%	71.60%	3.70%	0.00%	43.83%	23.46%	32.72%
24	6	≤16.7	<4	1447 (3.3)	94.26%	5.60%	0.14%	0.00%	72.36%	27.64%	0.00%
25	7b	>16.7	<4	445 (1.0)	32.36%	62.47%	4.72%	0.45%	40.90%	15.51%	43.60%
26	9-10	>16.7	10-<20	533 (1.2)	3.00%	67.73%	27.20%	2.06%	62.10%	0.00%	37.90%
27	8	>16.7	<4	289 (0.7)	18.69%	76.47%	3.46%	1.38%	27.68%	17.65%	54.67%
28	8	≤16.7	10-<20	125 (0.3)	16.80%	79.20%	4.00%	0.00%	22.40%	19.20%	58.40%
29	8	>16.7	4-<10	2032 (4.6)	18.85%	73.97%	6.69%	0.49%	23.47%	16.98%	59.55%
30	7b	≤16.7	≥20	53 (0.1)	16.98%	79.25%	3.77%	0.00%	43.40%	28.30%	28.30%
31	8	>16.7	≥20	424 (1.0)	3.77%	69.34%	24.76%	2.12%	28.30%	28.54%	43.16%
32	9-10	≤16.7	4-<10	141 (0.3)	9.93%	84.40%	4.96%	0.71%	57.45%	0.00%	42.55%
33	8	≤16.7	4-<10	490 (1.1)	26.73%	70.00%	3.06%	0.20%	24.90%	11.84%	63.27%
34	9-10	>16.7	<4	168 (0.4)	5.95%	74.40%	19.05%	0.60%	70.83%	0.00%	29.17%
35	9-10	≤16.7	10-<20	28 (0.06)	3.57%	82.14%	14.29%	0.00%	64.29%	0.00%	35.71%
36	8	≤16.7	<4	63 (0.14)	26.98%	69.84%	3.17%	0.00%	30.16%	14.29%	55.56%
37	7a	≤16.7	≥20	76 (0.2)	46.05%	46.05%	7.89%	0.00%	46.05%	39.47%	14.47%
38	9-10	≤16.7	≥20	14 (0.03)	7.14%	78.57%	7.14%	7.14%	57.14%	0.00%	42.86%
39	9-10	≤16.7	<4	20 (0.05)	30.00%	65.00%	5.00%	0.00%	45.00%	0.00%	55.00%

PSA: prostate-specific antigen.
 Level 1: pT2 and Gleason score 7a; Level 2: pT3/4 or Gleason score >7a and pN0cM0; Level 3: pN1cM0; Level 4: cM1.

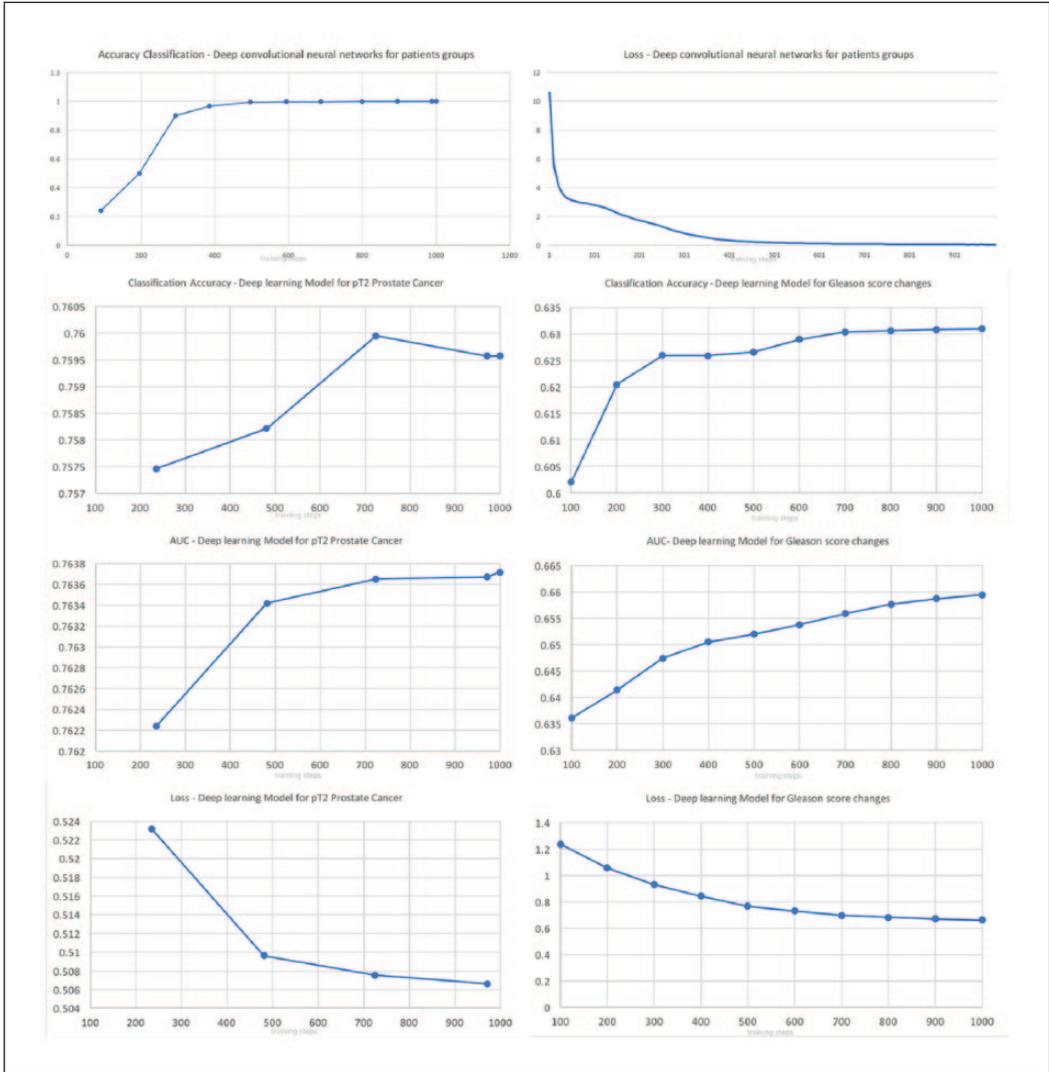


Figure 3. Diagrams illustrating the classification accuracy (CA), area under the curve (AUC), loss function during training episodes (training steps) for deep models for classification of organ-confined prostate cancer or identification of cases with Gleason score change. CA and loss function for the combination and deep model are shown in the first row of the diagrams.

classifying patients for treatment planning, for instance, the AJCC TNM staging system;²⁵ the D’Amico²⁶ classification, which categorized PSA to <4, 4–10, 10–20, and >20 ng/mL; or the recommended Gleason grouping (6, 7a, 7b, 8, 9–10).²⁷ Furthermore, we could determine a cutoff for the ratio of positive biopsy cores in our data sets. This cutoff of 16.7 percent reflected the number of positive cores defined by Epstein criteria for insignificant PCA in our data set (two positive cores)²⁸ and was used for the selection of patients for active surveillance.^{29,30} We identified similar benefits of categorization of input parameters to simplify the construction of the proposed model. However, the definition of thresholds for input parameters should be carefully considered and

Table 4. The model performances for prediction of pT2 prostate cancer and Gleason score changes.

Method	AUC	CA	FI score	Precision	Recall
By organ-confined PCa (pT2)					
Supportive vector machine	0.573	0.718	0.648	0.591	0.718
Adaptive boosting	0.715	0.748	0.737	0.726	0.748
Logistic regression	0.763	0.761	0.752	0.743	0.761
Classification tree	0.735	0.751	0.740	0.730	0.751
Naïve Bayes	0.762	0.757	0.748	0.740	0.757
Random forest classification	0.731	0.749	0.737	0.726	0.749
k-nearest neighbors	0.686	0.725	0.714	0.703	0.725
Dense neural network	0.762	0.759	0.850	0.772	0.946
Wide and dense neural network	0.762	0.760	0.849	0.772	0.943
By Gleason score upgrading					
Supportive vector machine	0.650	0.605	0.497	0.414	0.621
Adaptive boosting	0.630	0.613	0.526	0.475	0.588
Logistic regression	0.653	0.628	0.532	0.469	0.614
Classification tree	n.c.	n.c.	n.c.	n.c.	n.c.
Naïve Bayes	0.652	0.622	0.497	0.414	0.622
Random forest classification	0.636	0.612	0.530	0.484	0.585
k-nearest neighbors	0.586	0.573	0.521	0.516	0.527
Dense neural network	0.660	0.629	0.574	0.596	0.554
Wide and dense neural network	0.659	0.629	0.563	0.600	0.530

n.c.: cannot be calculated; AUC: area under the curve; CA: classification accuracy; PSA: prostate-specific antigen.

The bold numbers represent the top achieved results; Other results are close to the top results are also marked with bold.

should be clinically meaningful. Some approaches have been introduced, including Youden index, median, percentile, and selection of cutoff with high sensitivity or high specificity or optimal AUC.^{31,32} We emphasize first to apply well-accepted cutoffs of input parameters for reproducibility and to avoid misinterpretation of results from the model. When no well-accepted cutoffs were found in the literature, then using abovementioned statistical approaches can be helpful to determine the cutoff.

We applied the parameter selection to build the combination model by weighing the clinical meaning and association between the input and outcome data. Our approach enables further the extension of the combination possibility and deep learning model (CDLM) by weighing the information sources according to their importance (e.g. age or race have a lesser information weight than Gleason score or PSA), since the predictive value of age at diagnosis and race for advanced tumor and GIs upgrading is depending on GIs scores as shown by recent studies.^{33–35}

We preferred the neural network over conditional algorithms for the classification system developed from the parameter combination due to the high scalability and easier extensibility of the neural network. The depth of input parameters (subcategories) should be considered when developing a combination and deep neural network model. The depth of input parameters defines the number of possible combinations of input parameters. However, these combinations and outcome must be clinically realistic. For instance, the presence of lymph node metastases is unusual in patients with organ-confined PCas having GIs 6 and PSA levels below 10.^{36,37}

The possible combination reflects different clinical scenarios observed in clinical routine. The clinical outcome is a consequence of the combination of different pathologic features seen in patients. For instance, only four (0.08%) of the patients having preoperative Gleason score 6, a

percentage of positive cores ≤ 16.67 percent and a PSA level between 4 and <10 ng/mL had lymph node metastases; when the percentage of positive cores is >16.7 percent, the frequency of cases with lymph node metastases increases to 0.35 percent. When only the preoperative PSA level is changed to $10\text{--}<20$ ng/mL, the frequency of lymph node metastases increases to 1 percent. When both features (i.e. PSA: $10\text{--}<20$ ng/mL and $>16.7\%$ positive cores) are altered, the frequency of lymph node metastases increases to 1.77 percent. When the PSA level is changed to a level >20 ng/mL, the frequency of lymph node metastases increases to 4.98 percent. This observation explains one of the reasons of considering PSA >10 ng/mL and setting the maximal number of positive cores to two positive cores as eligibility criteria for active surveillance in very low-risk patients and considering PSA levels >20 ng/mL as indication for metastases screening.²

Previous work and prediction models were mostly based on regression models or support vector machine.^{38–58} In contrast, our work introduced the prediction model, which has first utilized the deep neural network with two layers for the prediction model development for PCa.

Our model, which is based on multilayer dense neural network, can predict organ-confined PCa with AUC of 76.2 percent or F1 score 84.9 percent higher than the current Partin's nomogram (a well-known nomogram in PCa) for organ-confined PCa with 70.4 percent (AUC).⁵⁹ In conclusion, we believe that prediction models based on multilayer dense neural network can perform better than conventional machine learning approaches.

We found that the classification accuracy of the deep learning models is equal to the highest classification accuracy of other machine learning approaches as given in Table 4. When we focus on the evaluation parameters, the deep learning models showed the best results in classification performance measured by F1 score, recall, and precision, implicating the possible strength of the deep learning model in the prediction of cancer outcomes. Table 5 lists some previous models with the classification accuracy performance and their methods.^{6–10} Due to the fact that all previous models have been tested on different data sets and it is insufficient to conduct a comparison analysis based on the results from these papers, we explicitly avoided any comparison analysis with previous models. In our opinion, it is essential to provide a validation set that can be used for comparison analysis between different models. However, this problem remains outside the scope of this article that evaluated the performance of different machine learning methods for the outcome prediction in PCa.

The inclusion of different data sources is essential to building an accurate deep model that can identify all possible combinations of clinical parameters. Our results show that a single institutional data set cannot cover all possible combinations of featured parameters. Therefore, more than one data set or national cancer registries are required to complete or verify all possible combinations.

In summary, this study utilized for the first time the multilayer dense neural network in predicting pathologic outcomes for PCa and introduced a novel model called "combination and deep model" that allows identifying subsets of patients and corresponding observed relative frequencies. Moreover, we confirm through comparison analyses the accuracy of deep learning approach in predicting outcome using the preoperative parameters to predict the outcome. We underline the importance of clinical knowledge for developing the deep learning approach and validating the results of the deep learning models. We recommend utilizing more than one data set to train the deep models that cover all possible combinations of clinical features to predict outcomes and to provide clinically reasonable results. We included data resources representing different data acquisition scenarios and integrated them into our models after data preparation. We emphasize further that the selection of clinical features for the deep models should be available in clinical routine and must have a clinical implication. Our combination and deep model is a helpful decision-aided tool for urologists to optimize the treatment strategy by classifying the patients into different groups

Table 5. An overview of the previous prediction models for organ-confined prostate cancer and Gleason score upgrading and their performance.

Organ-confined prostate cancer				
Previous studies	Variables	Cohort size	Statistical approach	Classification accuracy
Partin et al. ⁵	Biopsy Gleason sum, clinical stage, preoperative PSA	4133	Logistic regression analysis with the likelihood ratio chi-square test	72%
Kattan et al. ⁸	Preoperative PSA, clinical stage, primary and secondary biopsy Gleason sum, TRUS volume, millimeter core with cancer, millimeter core without cancer	409	Logistic regression	79%
Veltri et al. ¹⁰ and Haese et al. ⁷	Age at diagnosis, preoperative PSA, no. of cores positive, highest Gleason score, average % tumor involvement per core, presence of Gleason pattern 4/5, midcore with >5% tumor, base and/or midcore with >5% tumor	1287	Ordinal logistic regression and genetically engineered neural networks	94.9%
Gleason score upgrading				
Chun et al. ⁶	PSA, clinical stage, biopsy Gleason sum	4789	Logistic regression coefficients were used to develop and validate a nomogram	76%

PSA: prostate-specific antigen.

representing different risk constellations. Our novel model can be implemented in clinical routine by a smartphone app as a decision-aided tool for risk assessment for PCa. However, further study will be necessary to evaluate the acceptance of CDLM by physicians.

This study inherits some limitations that warrant mentioning. First, the pathologic evaluation was made by several pathologists and inherits the inter- and intra-observer variation. Second, there was a lack of information regarding patient comorbidities and the use of additional treatments (including radiation, systemic, salvage, and hormone therapies). Furthermore, documentation errors or misdiagnosis of metastatic disease may exist in SEER. However, SEER is the only comprehensive population-based database in the United States and represents an ideal approach to study the tumor progression in a large population with PCa, especially in recent time periods. Another limitation is that our classification model is focused only on patients who were treated by RPE. However, using these cases enabled the comparison between preoperative and final pathologic conditions.

Future work

Our future work will be focused on developing a complete model covering most of the preoperative parameters (e.g. magnetic resonance imaging (MRI) of the prostate and histology imaging) to improve our classification models for final pathologic outcomes in PCa. Moreover, we aim to improve the current risk classification system for PCa with the help of the combination and deep models.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: O.E. is supported by Dr Werner Jack Staedt Foundation, DOD.

Supplemental material

Supplemental material for this article is available online.

ORCID iDs

Okyaz Eminaga  <https://orcid.org/0000-0003-0861-3138>

Bernhard Breil  <https://orcid.org/0000-0002-2891-5558>

References

1. Jemal A, Siegel R, Xu J, et al. Cancer statistics, 2010. *CA Cancer J Clin* 2010; 60: 277–300.
2. Heidenreich A, Bellmunt J, Bolla M, et al. EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and treatment of clinically localised disease. *Eur Urol* 2011; 59: 61–71.
3. Lughezzani G, Briganti A, Karakiewicz PI, et al. Predictive and prognostic models in radical prostatectomy candidates: a critical analysis of the literature. *Eur Urol* 2010; 58: 687–700.
4. D’Amico AV, Renshaw AA, Arsenault L, et al. Clinical predictors of upgrading to Gleason grade 4 or 5 disease at radical prostatectomy: potential implications for patient selection for radiation and androgen suppression therapy. *Int J Radiat Oncol Biol Phys* 1999; 45: 841–846.
5. Partin AW, Mangold LA, Lamm DM, et al. Contemporary update of prostate cancer staging nomograms (Partin Tables) for the new millennium. *Urology* 2001; 58: 843–848.
6. Chun FK, Briganti A, Shariat SF, et al. Significant upgrading affects a third of men diagnosed with prostate cancer: predictive nomogram and internal validation. *BJU Int* 2006; 98: 329–334.
7. Haese A, Chaudhari M, Miller MC, et al. Quantitative biopsy pathology for the prediction of pathologically organ-confined prostate carcinoma: a multiinstitutional validation study. *Cancer* 2003; 97: 969–978.
8. Kattan MW, Eastham JA, Wheeler TM, et al. Counseling men with prostate cancer: a nomogram for predicting the presence of small, moderately differentiated, confined tumors. *J Urol* 2003; 170: 1792–1797.
9. Partin AW, Yoo J, Carter HB, et al. The use of prostate specific antigen, clinical stage and Gleason score to predict pathological stage in men with localized prostate cancer. *J Urol* 1993; 150: 110–114.
10. Veltri RW, Miller MC, Partin AW, et al. Prediction of prostate carcinoma stage by quantitative biopsy pathology. *Cancer* 2001; 91: 2322–2328.
11. Regnier-Coudert O, McCall J, Lothian R, et al. Machine learning for improved pathological staging of prostate cancer: a performance comparison on a range of classifiers. *Artif Intell Med* 2012; 55: 25–35.
12. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017; 7: 16878.
13. Auffenberg GB, Ghani KR, Ramani S, et al. 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.
14. Herzberg S, Rahbar K, Stegger L, et al. Concept and implementation of a single source information system in nuclear medicine for myocardial scintigraphy (SPECT-CT data). *Appl Clin Inform* 2010; 1: 50–67.
15. Breil B, Semjonow A, Muller-Tidow C, et al. HIS-based Kaplan-Meier plots—a single source approach for documenting and reusing routine survival information. *BMC Med Inform Decis Mak* 2011; 11: 11.

16. Epstein JI, Allsbrook WC Jr, Amin MB, et al. The 2005 international society of urological pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. *Am J Surg Pathol* 2005; 29: 1228–1242.
17. Eminaga O, Hinkelammert R, Abbas M, et al. Prostate cancers detected on repeat prostate biopsies show spatial distributions that differ from those detected on the initial biopsies. *BJU Int* 2015; 116: 57–64.
18. Breil B, Semjonow A and Dugas M. HIS-based electronic documentation can significantly reduce the time from biopsy to final report for prostate tumours and supports quality management as well as clinical research. *BMC Med Inform Decis Mak* 2009; 9: 5.
19. Eminaga O, Abbas M, Hinkelammert R, et al. CMDX©-based single source information system for simplified quality management and clinical research in prostate cancer. *BMC Med Inform Decis Mak* 2012; 12: 141.
20. Edge SB. *American Joint Committee on Cancer: AJCC cancer staging handbook: From the AJCC cancer staging manual*. 7th ed. New York: Springer, 2010.
21. Cheng H-T, Koc L, Harmsen J, et al. Wide & deep learning for recommender systems. In: Proceedings of the 1st workshop on deep learning for recommender systems, Boston, MA, 15 September 2016.
22. Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 2014; 15: 1929–1958.
23. Kingma DP and Ba J. Adam: a method for stochastic optimization. Paper Presented at the 3rd international conference for learning representations, San Diego, CA, 22 December, 2014.
24. LeCun Y, Bengio Y and Hinton G. Deep learning. *Nature* 2015; 521: 436–444.
25. Compton CC, Byrd DR, Garcia-Aguilar J, et al. *AJCC cancer staging atlas: a companion to the seventh editions of the AJCC cancer staging manual and handbook*. 2nd ed. New York: Springer, 2012.
26. D’Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* 1998; 280: 969–974.
27. Epstein JI, Zelefsky MJ, Sjoberg DD, et al. A contemporary prostate cancer grading system: a validated alternative to the Gleason score. *Eur Urol* 2016; 69: 428–435.
28. Epstein JI, Walsh PC, Carmichael M, et al. Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer. *JAMA* 1994; 271: 368–374.
29. Vellekoop A, Loeb S, Folkvaljon Y, et al. Population based study of predictors of adverse pathology among candidates for active surveillance with Gleason 6 prostate cancer. *J Urol* 2014; 191: 350–357.
30. Bul M, Zhu X, Valdagni R, et al. Active surveillance for low-risk prostate cancer worldwide: the PRIAS study. *Eur Urol* 2013; 63: 597–603.
31. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950; 3: 32–35.
32. Leeflang MM, Moons KG, Reitsma JB, et al. Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: mechanisms, magnitude, and solutions. *Clin Chem* 2008; 54: 729–737.
33. Caster JM, Falchook AD, Hendrix LH, et al. Risk of pathologic upgrading or locally advanced disease in early prostate cancer patients based on biopsy Gleason score and PSA: a population-based study of modern patients. *Int J Radiat Oncol Biol Phys* 2015; 92: 244–251.
34. Dinh KT, Mahal BA, Ziehr DR, et al. Incidence and predictors of upgrading and up staging among 10,000 contemporary patients with low risk prostate cancer. *J Urol* 2015; 194: 343–349.
35. Eminaga O, Hinkelammert R, Titze U, et al. The presence of positive surgical margins in patients with organ-confined prostate cancer results in biochemical recurrence at a similar rate to that in patients with extracapsular extension and PSA ≤ 10 ng/ml. *Urol Oncol* 2014; 32: 32.e17–25.
36. Liu JJ, Lichtensztajn DY, Gomez SL, et al. Nationwide prevalence of lymph node metastases in Gleason score 3 + 3 = 6 prostate cancer. *Pathology* 2014; 46: 306–310.
37. Weckermann D, Goppelt M, Dorn R, et al. Incidence of positive pelvic lymph nodes in patients with prostate cancer, a prostate-specific antigen (PSA) level of $< \text{or} = 10$ ng/mL and biopsy Gleason score of $< \text{or} = 6$, and their influence on PSA progression-free survival after radical prostatectomy. *BJU Int* 2006; 97: 1173–1178.
38. Wang J, Wu CJ, Bao ML, et al. 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.

39. Kim SH, Kim S, Joung JY, et al. Lifestyle risk prediction model for prostate cancer in a Korean population. *Cancer Res Treat* 2018; 50: 1194–1202.
40. Xiao J, Wang S, Shang J, et al. DWCoX: a density-weighted Cox model for outlier-robust prediction of prostate cancer survival. *F1000Res* 2016; 5: 2806.
41. Wang Q, Li YF, Jiang J, et al. The establishment and evaluation of a new model for the prediction of prostate cancer. *Medicine (Baltimore)* 2017; 96: e6138.
42. Guinney J, Wang T, Laajala TD, et al. Prediction of overall survival for patients with metastatic castration-resistant prostate cancer: development of a prognostic model through a crowdsourced challenge with open clinical trial data. *Lancet Oncol* 2017; 18: 132–142.
43. Coley RY, Fisher AJ, Mamawala M, et al. A Bayesian hierarchical model for prediction of latent health states from multiple data sources with application to active surveillance of prostate cancer. *Biometrics* 2017; 73: 625–634.
44. Cosma G, Acampora G, Brown D, et al. Prediction of pathological stage in patients with prostate cancer: a neuro-fuzzy model. *PLoS ONE* 2016; 11: e0155856.
45. Peters M, van der Voort van Zyp JR, Moerland MA, et al. Development and internal validation of a multivariable prediction model for biochemical failure after whole-gland salvage iodine-125 prostate brachytherapy for recurrent prostate cancer. *Brachytherapy* 2016; 15: 296–305.
46. Peng H, Zhao W, Tan H, et al. Prediction of treatment efficacy for prostate cancer using a mathematical model. *Sci Rep* 2016; 6: 21599.
47. Murray NP, Aedo S, Reyes E, et al. Prediction model for early biochemical recurrence after radical prostatectomy based on the Cancer of the Prostate Risk Assessment score and the presence of secondary circulating prostate cells. *BJU Int* 2016; 118: 556–562.
48. Kerkmeijer LG, Monninkhof EM, van Oort IM, et al. PREDICT: model for prediction of survival in localized prostate cancer. *World J Urol* 2016; 34: 789–795.
49. Chanrion MA, Sauerwein W, Jelen U, et al. The influence of the local effect model parameters on the prediction of the tumor control probability for prostate cancer. *Phys Med Biol* 2014; 59: 3019–3040.
50. Akamatsu S, Takahashi A, Takata R, et al. Reproducibility, performance, and clinical utility of a genetic risk prediction model for prostate cancer in Japanese. *PLoS ONE* 2012; 7: e46454.
51. Chen SH, Ip EH, Xu J, et al. Using graded response model for the prediction of prostate cancer risk. *Hum Genet* 2012; 131: 1327–1336.
52. Williams SB, Salami S, Regan MM, et al. Selective detection of histologically aggressive prostate cancer: an Early Detection Research Network Prediction model to reduce unnecessary prostate biopsies with validation in the Prostate Cancer Prevention Trial. *Cancer* 2012; 118: 2651–2658.
53. Kattan MW and Gerds TA. Making and evaluating a statistical prediction model for the absolute risk of prostate cancer recurrence. *Cancer* 2011; 117: 5026–5028.
54. Hosmer W, Malin J and Wong M. Development and validation of a prediction model for the risk of developing febrile neutropenia in the first cycle of chemotherapy among elderly patients with breast, lung, colorectal, and prostate cancer. *Support Care Cancer* 2011; 19: 333–341.
55. Buffa FM, Flux GD, Guy MJ, et al. A model-based method for the prediction of whole-body absorbed dose and bone marrow toxicity for ¹⁸⁶Re-HEDP treatment of skeletal metastases from prostate cancer. *Eur J Nucl Med Mol Imaging* 2003; 30: 1114–1124.
56. Cohen RJ, Chan WC, Edgar SG, et al. Prediction of pathological stage and clinical outcome in prostate cancer: an improved pre-operative model incorporating biopsy-determined intraductal carcinoma. *Br J Urol* 1998; 81: 413–418.
57. Carter HB, Partin AW and Coffey DS. Prediction of metastatic potential in an animal model of prostate cancer: flow cytometric quantification of cell surface charge. *J Urol* 1989; 142: 1338–1341.
58. Tosoian JJ, Chappidi M, Feng Z, et al. Prediction of pathological stage based on clinical stage, serum prostate-specific antigen, and biopsy Gleason score: Partin Tables in the contemporary era. *BJU Int* 2017; 119: 676–683.
59. Leyh-Bannurah SR, Gazdovich S, Budaus L, et al. Population-based external validation of the updated 2012 Partin Tables in contemporary North American Prostate Cancer Patients. *Prostate* 2017; 77: 105–113.