

Supplementary Table 1. General characteristics of included studies

#	Author	Year	County	Machine learning method	Software	Source of data	Sample size	Train/ test	Model performance	Prospective /Retrospective
1	Ping Ge [23]	2015	China	•LR •ANN	•MATLAB	Hospital	586	Train: 2/3 of data Test: 1/3 of data	Logistic regression: Sensitivity: 86.23%, Specificity:86.16%, Accuracy: 86.18%, Youden Index: 72.39% Neural network: Sensitivity: 85.63 %, Specificity: 88.54%, Accuracy: 87.71%, Youden Index: 74.17%	Retrospective
2	Li-Hong Xiao [24]	2016	China	•RF	•R software	Hospital	941	Not mentioned	Sensitivity: 65.64%, Specificity: 93.83%, Accuracy:83.10%, Positive predictive value: 86.72%, Negative predictive value: 81.64%	Retrospective
3	Nguyen Thi Hong Nhung [25]	2016	Vietnam	•SVM	•R software	Hospital	1110	Normal (train:544, test:231) BPH (train:187, test:76) Pca (train:48, test:23)	Sensitivity:95.1%, Specificity:84.6%	Retrospective
4	Yi-Yan Zhang [37]	2017	China	•DT	•R software	Hospital	392	Train: 240 Test: 152	Sensitivity: 86.6%, Specificity: 78.1%	Retrospective
5	Farzad Allameh [26]	2017	Iran	•LR •MLP	•SAS • MATLAB	Hospital	566	Not mentioned	Logistic regression: Sensitivity: 81.2%, Specificity: 69.1%, AUC:0.885%, Correct percentage: 81.2%	Retrospective

									Neural network: Sensitivity: 83.9%, Specificity: 71.3%, AUC: 0.901, Correct percentage: 85.3 %	
6	Sam Li-Sheng Chen [34]	2018	Taiwan	•Bayesian predictive model	•SAS •Winbugs	Hospital	5269	Not mentioned	AUC increased from 77.0% based on PSA alone to 83.8% based on PSA combined with 7 SNPs whereas 80.6% of the AUC was noted when PSA was combined with 66 SNPs	Prospective
7	Juan Enrique Sánchez Lasheras [40]	2018	Spain	•MARS •ANN •SVM	•Not mentioned	MCC-Spain study database	1912	Train: 80% Validation: 20%	ANNs: AUC of 0.62006, MARS: AUC of 0.569312, SVM: AUC of 0.65212	Retrospective
8	Damjan N Pantic [38]	2018	Serbia	•CART	•SPSS	Hospital	97	Not mentioned	Accuracy: 78.3%, Specificity: 80.3%, AUC: 82.6%, Sensitivity: 73.1%	Retrospective
9	Miroslav M. Stojadinović [39]	2018	Serbia	•CART •LR	•SPSS	Clinical Centre	221	Not mentioned	Sensitivity: 92.4%, AUC:83.3%, Accuracy: 78.7%	Retrospective
10	Juan Enrique Sánchez Lasheras [41]	2018	Spain	•Recursive partitioning •MARS	•Not mentioned	MCC-Spain study database	1912	Not mentioned	Not clear	Retrospective
11	Heidi A. Hanson [35]	2019	United States	•RF	•R software	SEER registry, HHS, AHRF, CDC, AHRQ	514, 878	Not mentioned	Not mentioned	Retrospective
12	Juan Enrique	2019	Spain	•A hybrid algorithm	•Not mentioned	MCC-Spain	2605	Train: 80% Test:20%	Not mentioned	Retrospective

	Sánchez Lasheras [27]			based on support vector machines and genetic algorithm		study database				
13	Maurizia Mello-Grand [36]	2019	Italy	•Linear models with empirical Bayesian approach •LR	•R software	Hospital	242	Not mentioned	Sensitivity: 87%, Specificity:35%, Accuracy:68%	Prospective
14	Jue Tao Lim [43]	2019	Singapore	•Linear probit regression models •Bayesian kernel machine regression •Generalized additive models	•R software	SPCS study database	255	Not mentioned	Not mentioned	Retrospective
15	Henry Barlow [28]	2019	Australia	•KNN •SVM •DT •RF •MLP •ADA •QDA	•Scikit-learn library (Python)	PLCO dataset	1130	Train: 75% Test: 25%	10-fold cross validation accuracy: KN: 0.876, SVM2: 0.894, QD: 0.919, DT: 0.838, RF: 0.835, MLPC: 0.845, ADA: 0.843	Retrospective
16	Rachel Dankner [44]	2019	Israel	•Cox regression	•Not mentioned	Electronic records from the health mainten	315890	Not mentioned	Not mentioned	Retrospective

						ance organiza tion in Israel, Clalit Health Services				
1 7	Marlon Perera [29]	2020	Germany	•ANN	•Not mentioned	Local institution and PLCO database	4548	Train: 3638 Test: 910	Sensitivity: 80%, Specificity: 45.3%,	Retrospective
1 8	Mingqi Li [30]	2020	China	•ADA •RF	•Scikit-learn library (Python)	Prostate Cancer Dataset of the National Center for Clinical Medical Science Data	3000	Train:70% Test:30%	AdaBoost: Accuracy: 0.88 , Recall:0.95, f1-score: 0.91, AUC: 0.98 Random forest: Accuracy: 0.89, Recall: 0.97, f1-score:0.93 , AUC:0.99	Retrospective
1 9	Simon P Hood [45]	2020	UK	•Ensemble Subspace KNN	•Kaluza	Hospital	71	Train: 64 Validation:7	Across 30 runs, the average performance was an AUC of 81.8%, an ORP TPR of 83.6%, an FPR of 20.1%, and an accuracy of 82.1%.	Retrospective
2 0	Shan-Ju Yeh [42]	2022	Taiwan	•DNN	•Python	National Center for Biotechnology Information	202	Train: 70% Test: 30%	Not mentioned	Retrospective

(NCBI) dataset										
2 1	Xinru Zhang [31]	2023	China	•KNN •SVM •DT •RF •NB •Gradient booster tree •LR	•MATLAB	Hospital	174	Train: 70% Test: 30%	Best accuracy: 73.65%	Retrospective
2 2	Maoliang Zhang [32]	2023	China	•ANN •LR •SVM •DT •RF •KNN	•Python	Hospital	301	Train: 240 Test: 61	Sensitivity: 80%, Specificity: 80%, AUC: 0.855	Retrospective
2 3	Muyu Wu [33]	2024	China	•SVM	•Python	Hospital	1203	Train: 960 Test: 240	Accuracy: Normal: 83.3, BPH: .91.6, Pca: 83.3	Retrospective

Abbreviations: LR: Logistic Regression; ANN: Artificial Neural Network; RF: Random Forest; DNN: Deep Neural Network; SVM: Support Vector Machine; DT: Decision Tree; KNN: K-Nearest Neighbors; NB: Naive Bayes; QDA: Quadratic Discriminant Analysis; MLP: Multilayer Perceptron; CART: Classification And Regression Tree; Pca: Prostate cancer; BPH: Benign Prostatic Hyperplasia; AUC: Area Under the Curve; MARS: Multivariate Adaptive Regression Splines

