CLINICAL RISK FACTORS AND RADIOLOGICAL FINDINGS USING MACHINE LEARNING FOR PATIENTS WITH PIRADS 3 CATEGORY LESIONS.

PORRECA ANNAMARIA[1], MARCHIONI MICHELE[2], BORRELLI PAOLA[1], DI NICOLA MARTA[1]

[1]Laboratory of Biostatistics, Department of Medical, Oral and Biotechnologies Sciences, "G. d'Annunzio" University, Chieti-Pescara, 66100 Chieti, Italy

[2]Urology Unit, Department of Medical, Oral and Biotechnological Sciences, G. D'Annunzio University Chieti-Pescara, 66100 Chieti, Italy

Introduction

The Prostate Imaging–Reporting and Data System (PIRADS) is a framework used to evaluate prostate cancer risk, ranging from category 1 (no risk) to category 5 (high risk). PIRADS category 3 represents an uncertain risk level. The PRECISION trial showed that while multiparametric MRI (mpMRI) is useful for detecting cancer-based on PIRADS categories, it does not incorporate clinical data into the risk assessment. To enhance clinically significant prostate cancer assessments, combining clinical risk factors with mpMRI data could improve the detection rate, especially for PIRADS category 3 lesions [1,2]. Recent advances in data science and machine learning (ML) have enabled the development of sophisticated medical analytic systems. ML models can effectively weigh various risk factors, making them highly relevant in the biomedical field [3]. This study aimed to create ML models to assist in clinical decisions regarding whether to perform a biopsy immediately or monitor the patient instead. The goal was to develop an ML model that assesses prostate cancer risk, potentially reducing unnecessary biopsies for patients with a low probability of cancer. This approach would optimise hospital resources such as surgical teams, nurse specialists, operating rooms, and hospital beds. This retrospective, single-centre study involved 2209 patients with at least one PIRADS category 3 who underwent magnetic resonance imaging/ultrasonography (MRI/US) fusion biopsies.

Objective

Aiming to improve the low cancer detection rates in patients with PIRADS 3 category lesions, researchers developed machine learning (ML) models to help decide whether to conduct prostate biopsies or monitor clinical data without biopsy results.

Methods

The original dataset was split into training and testing using 80% of the samples to make up
the training and the remaining part (20%) of the test dataset. The ML models, designed as supervised classification problems with binary outputs (cancerous or benign tissue), used clinical inputs such as age, PSA density levels, the number of previous biopsies, digital rectal examinations (DRE), maximum diameter of all observable PIRADS 3 category lesions. Radiology inputs included the number of lesions, maximum lesion diameter, location, and lesion zone. Less significant inputs were subsequently removed. Variables were selected based on the existing literature by adding irrelevant variables to the classifications [1]. The study applied various ML algorithms, including Logistic Regression, Support Vector Machine, Naive Bayes, Decision Tree, Random Forest, and eXtreme Gradient Boosting Tree (XGBoost). The performance of these models was assessed using the Area Under Curve (AUC) and 95% Confidence Interval (CI), Sensitivity, Specificity, Positive

and Negative Predicted Value. The mean SHAP (SHapley Additive exPlanations) scores were used for feature contribution to the models.

Results

Table 1. Machine learning models' performance.

Performance	LR	KNN	SVM	NB	DT	RF	XGBoost
AUC	**0.851**	**0.742**	**0.842**	**0.801**	**0.864**	**0.867**	**0.871**
Low-95%CI	**0.814**	**0.699**	**0.804**	**0.761**	**0.829**	**0.831**	**0.836**
High-95%CI	**0.883**	**0.782**	**0.874**	**0.837**	**0.895**	**0.897**	**0.901**
Sensitivity	0.941	0.735	0.982	0.973	0.936	0.904	0.945
Specificity	0.762	0.659	0.704	0.632	0.794	0.830	0.798
Pos Pred Value	0.795	0.679	0.765	0.722	0.817	0.839	0.821
Neg Pred Value	0.929	0.717	0.975	0.959	0.927	0.898	0.937

LR: Logistic Regression; KNN: K Nearest Neighbor; SVM: Support Vector Machine; DT: Decision Tree; RF: Random Forest; XGBoost: eXtreme Gradient Boosting Tree.

Table 1 reports the final models' performances. The best-performing model was XGBoost, with an AUC of 0.871 and 94.5% sensitivity. The second-ranking model was Random Forest, with an AUC of 0.867 and 90.4% sensitivity. The worst-performing model is K Nearest Neighbor, with an AUC of 0.782. Indeed, the mean SHAP scores in XGBoost list the variables by importance as follows: DRE, PSA density levels, maximum diameter, and age.

Conclusions

In this context, DRE is important because it provides valuable clinical information about the prostate. It allows healthcare providers to assess the prostate gland's size, shape, and texture, which can help identify abnormalities that might not be detected through imaging alone. DRE can reveal signs of prostate cancer, such as hard or irregular areas, and can complement other diagnostic tools like MRI and PSA testing. Incorporating DRE findings into the assessment model can enhance the accuracy of prostate cancer risk evaluation and decision-making regarding biopsies, particularly for patients with PIRADS 3 category lesions. The ML models demonstrated strong performance and showed promise for aiding decisions on whether to perform MRI/US fusion prostate biopsies for patients with PIRADS 3 category lesions at the study centre.

References

1. Aussavavirojekul, P.; Hoonlor, A.; Srinualnad, S. Optimization of Clinical Risk-Factor Interpretation and Radiological Findings with Machine Learning for PIRADS Category 3 Patients. The Prostate 2022; 82:235-244. doi:10.1002/pros.24266.
2. Hamilton, W.; Sharp, D.J.; Peters, T.J.; Round, A.P. Clinical Features of Prostate Cancer before Diagnosis: A Population-Based, Case-Control Study. Br J Gen Pract 2006; 56:756-762.
3. Vidanagamachchi, S.M.; Waidyarathna, T. Linear and Non-Linear Classifiers for Clinical Risk Factor Analysis of Cancer Patients. In Proceedings of the 2018 7th International Conference on Bioinformatics and Biomedical Science; ACM: Shenzhen, China, June 23 2018; pp. 78-81.