BIU Journal of Basic and Applied Sciences 10(2): 204 – 210, 2025.

ISSN: 2563-6424

MACHINE LEARNING MODELS FOR PREDICTION OF PROSTATE CANCER

***IZOGIE, L. E.,[1] AKAZUE, M. I.[2] AND IHAMA, E.I.[1]**
[1]Department of Computer Science and Information Technology, School of Applied Sciences, Edo State Polytechnic, Usen, Benin City, Nigeria.
[2]Department of Computer Sciences, Faculty of Science, Delta State University, Abraka, Nigeria.
*Corresponding author: eyoski@yahoo.com

ABSTRACT
Prostate cancer comprises of the enlargement or increase in the magnitude of the prostate biopsies and a shortage of urological pathologists, and this constrains the diagnosis of prostate cancers. There are several predictive model to predict and manage prostate cancer, developed by different authors which involves the interpretation of data using aggregate mode to assist in determination and making accurate decision. Different means sure as PSA rate, MRI guided biopsies, genomics biomarkers, and Gleason scaling system are utilized for predicting the risk, to stratify, and observe patients during individual follow-ups. However, identification tracking and successive risk categorization oftentimes alter itself to a epochal subjectivity. The use of machine leaning models via Artificial intelligence (AI) allows practitioner to accept challenging relationships and manage large data sets, which is a difficult task and is inordinately difficult and time overwhelming for humans. The use of AI algorithms helps to reduce the level of subjectivity by exploitation less resources while improving the general efficiency and quality in prostate cancer identification and management. This study discusses different models for prostate detection.

KEYWORDS: *Cancer identification, Artificial Intelligence, Risk categorization, Prostate cancer, Decision*

INTRODUCTION

This study discuses some machine learning model for prediction of prostate cancer. Cancer is deadly ailment among humans, but Prostate cancer is becoming one of the commonly diagnosed non-skin disease in men. It is proposed that in every six American men, one is likely to be affected by this sickness throughout their lifetime (Torre *et al*., 2012). Prostate cancer management comprises of enlarged size of prostate biopsies and a shortage of urological pathologists, this is a constrain on the diagnosis of prostate cancers (Strom *et al*., 2020). Furthermore, the occurrence of an unpredictability of pathological grouping can consequently result in an over treatment or under treatment of prostate cancer (Harmon *et al*., 2019).

In prostate cancer, the use of AI has shown to be beneficial and aid in standardizing pathological grading to assess prostate cancer stratification and treatment. Additionally, AI shows promise in automating the assessment of characterization and severity of prostate cancer based on image-based tasks, including histopathologic, MRI, and biomarker diagnosis (Harmon *et al*., 2019). Furthermore, certain patients that are diagnosed with prostate cancer which is thought to be more indolent can continue with repeated forms of surveillance including prostate biopsies, PSA, and other forms of digital testing through MRI or rectal examinations unless they experience any physiological side effects (Carroll and Mohler, 2018).

Artificial Intelligence (AI) and Machine Learning (ML) in healthcare is a novel area of study, which has lately gained serious attention. AI frequently uses artificial neural networks (ANN) that use numerical simulations that are directly stimulated by and partly modelled on biological neural networks. They are capable of modelling and processing nonlinear relationships between inputs and outputs in parallel. AI can help improve these forms of surveillance and will be amongst some of the essential tools to urological pathologists and to the field of urology as a whole as technology continues to improve and help patient prognosis over time. In this present study, we systematically reviewed the literature to determine how AI can be used to improve the diagnosis and management of prostate cancer.

Kumar *et al*. (2017) model was devoted to recurrence prediction by creating 2 different Convolutional Neural Networks (CNN) analysing H&E images. The first CNN detected individual nuclei while the second CNN classified the patches around the nuclear centres. There was a voting process which yielded probabilities of recurrence from the patients. The model was trained on 80 case/control pairs and then validated on 30 recurrent and 30 non-recurrent controls. The end yield was 0.81 AUC for accuracy. Therefore, creating a deep learning tool that considers a recurrence in combination with other risk factors to come up with a new scoring system that could be more accurate than currently used.

LITERATURE REVIEW

The procedure for the identification of eminent PSA level is viewed as one of the commonest medical instruments for identification of prostate cancer. Stojadinovic *et al*. (2018) viewed several reasons and determined that the PSA mass was the key pivotal variable and revealed the decision tree providing a mesh advantage compared to a logistic regression model. In this retrospective study, the model had a potential to reduce unnecessary biopsies without missing significant diagnoses. It is projected that this ailment is likely to affect one of every six American throughout their lifespan (Torre *et al*., 2012).

Ge *et al*. (2015) established a logistic regression and ANN model to aid in diagnosis of prostate cancer. They utilized predictors including age, percent free PSA, prostate volume, and PSA density from 586 men with prostate cancer confirmed by biopsy to train their algorithms. While they found no significant difference between the two models, they concluded that both models have high diagnostic validity and have potential to be included in practice to prevent unnecessary biopsies.

Artificial Neural Network and Histopathologic Diagnosis of Prostate Cancer

The histopathological identification of prostatic adenocarcinoma is frequently vital for creating analysis of prostate cancer Humphrey (2017). Nevertheless, occasionally histopathologic changes happen owing to inter- and intra-observer variability. This variability could impact the evaluation of biologic aggressiveness of prostate cancer and the identification of patients at high risk for progression (Waliszewski *et al.,* 2015).

Artificial intelligence has interfaced in accurately predicting prostate cancer, localize and grade histopathologic slides. Initial studies focused on detecting the difference in histology slides between Gleason score. Gleason grading allows the stratification of the aggressiveness of prostate cancer into low (grade group 1), intermediate (grade group 2 and 3) or high-risk group (grade group 4 and 5). Group 1 is designated by Gleason score (GS) 6, group 2 by GS7 or 3+4, group 3 by GS7 or 4+3, group 4 by GS8, and group 5 by GS9-10, respectively.

The authors focused on detecting the difference between GS3 and GS4 (Bhele *et al.,* 2014). In another study by Lawrentschuk *et al.* (2011), a polychotomous logistic regression (PR) model and an ANN was developed for predicting the biopsy results using 3025 men undergoing biopsy with PSA <10ng/dL.

Clinical predictors of age, PSA, abnormal digital rectal examination, positive transrectal ultrasound (TRUS) and prostate volume were also considered during the development. It was suggested that an inclusion of additional predictors may improve the performance of the models; however, the ANN was unable to distinguish between the four biopsy outcomes that were utilized in their validation.

Artificial Neural Network and Magnetic Resonance Imaging (MRI) Diagnosis of Prostate Cancer

Magnetic resonance imaging (MRI) has been investigated as a modality for prostate cancer detection and determination of aggressiveness. Moore *et al*. (2014).MR spectroscopic T2-weighted MR imaging, and apparent diffusion coefficient (ADC) have been valuable tools to assess prostate cancer, but their use still lacks consensus (Fehr *et al*., 2015). Expert users can easily detect malignant tumours, but determining their aggressiveness using MRI with good reliability is more difficult. Thus, automatic classification using machine learning has been proposed as a solution to provide more accurate and consistent results to aid clinicians' management. In a similar context, Fehr *et al*. (2015) proposed a technique of utilizing machine learning-based automatic classification of prostate cancer aggressiveness by combining apparent diffusion coefficient (ADC) and T2-weighted MRI-based texture features. This technique was able to differentiate between the high GS (>7) and low GS6, as well as between 7 (3+4) and 7 (4+3) cancers. Furthermore, this model could distinguish between high and low Gleason grades with 93% accuracy for cancers occurring in both peripheral (PZ) and transition (TZ) zones and 92% for cancers occurring in the PZ alone. These results were far superior to using only ADC and suggest that this method can help provide reasonably accurate classification of Gleason patterns.

Antonelli *et al*. (2019) in their study used quantitative MRI and clinical features from 164 men to construct machine learning classifiers. Following model validation, these classifiers were able to predict Gleason 4 in prostate tumours with greater accuracy than the three board-certified radiologists which participated in the study. It was suggested that these AI classifier tools could be useful to non-invasively detect the progression of tumours and aid in decisions regarding active surveillance programs.

Toivonen *et al*. (2019) developed machine learning tools to predict prostate cancer aggressiveness by using optimized high-quality MRI data sets. A classifier system was created based on multiple texture features of high-quality T2 weighted images, DWI (diffusion weighted imaging), and T2 relaxation maps from 100 patients for prediction of PCa Gleason score dichotomized as 3+3 (low risk) vs >3+3 (high risk). Results suggested that texture feature analysis of DWI, post-processed using mono-exponential and kurtosis models, and T2w demonstrated good classification performance for Gleason score of prostate cancer. Together with these multiple studies which reflect the potential of AI ANNs to allow more effective integration with modern surveillance tools such as MRI to effectively help in prostate cancer surveillance.

Artificial Neural Network in Biomarker Diagnosis and Risk Stratification

Diagnosis and prognosis of prostate cancer is guided by PSA level testing. However, concerns about absolute accuracy can sometimes lead to patients being given invasive treatment options when active surveillance might provide better outcomes in these men Choyke (2017). Over the past 10 years, avalanches of biomarkers have been identified and are put into clinical assays (Jin *et al*., 2020). While several of these biomarkers have been studied and characterized, there is no standard overlap between all of these assays due to the function of each assay, and there is no absolutely perfect list of biomarkers to look at when predicting diagnosis and prognosis. Thus, it is important to be able to identify and evaluate any new biomarkers for their clinical significance in a way that is both meaningful and accurate. Therefore, ANNs can play an instrumental role in analysing and validating the biomarkers. For example, one study suggested that Ki67 is an important marker of survival and disease progression (Fruge *et al.,* 2020). Green *et al*. (2016) created an ANN that was designed to validate Ki67 gene expression while comparing it to another potential candidate in DLX2. Univariate analysis showed that both Ki67 and DLX2 were significant in predictiveness of future metastases. Nevertheless, only 6.8% of prostate cancer patients have high expression of Ki67. Thus, this study showed that these 2 biomarkers could be used to identify candidates for targeted therapy only (Green *et al*., 2016).

In addition to gene expression, proteomics can also be useful when identifying potential biomarkers. For example, Kim et al. developed a novel approach by combining targeted proteomics with computational biology to discover new potential proteomic signatures for prostate cancer. The study started with 133 differentially expressed proteins that were evaluated with synthetic peptides in a 74-patient cohort. Then they applied machine learning approaches to develop clinical predictive models using these candidates. Results showed that computationally guided proteomics can be used to discover novel non-invasive biomarkers. Overall, there are numerous studies which reflect the potential of AI ANNs to allow more effective identification and validation of biomarkers to aid in prostate cancer surveillance (Green *et al*., 2016).

Patients diagnosed with prostate cancer often remain confused due to treatment options that are available to them. Thus, understanding how certain therapies are put into place may allow for ease of mind and improvements in patient satisfaction. Auffenberg *et al*. (2019) developed a registry that used ANN to better allow patients to take charge of their care. This registry, which is called askMUSIC, takes data from 45 urology practices within the Michigan Urological Surgery Improvement Collaborative (MUSIC). This registry data is used to create a random forest machine learning model which could predict prostate cancer treatment options. Patients can go to askMUSIC website and interact with the registry data and predicted treatment to show therapy options to alleviate fear about a given therapy from the patient perspective (Auffenberg *et al*., 2019).

Alitto *et al*. (2017) developed a PRODIGE project; it used an Umbrella Protocol that focused on standardization of data sharing. Within this protocol, a standardized knowledge sharing process is implemented using semi-formal ontology representing clinical variables. This process can be adapted to use with machine learning or traditional statistics. The standardization of these techniques supports the multifactorial decision support systems (DSS) which can be seen as the basis for future patient-level support therapy decisions. Together, there are multiple studies which reflect the potential of AI ANNs to allow the development of effective patient-centric tools to help with educating the patients about the treatment options and disease progression/regression.

CONCLUSION

The use of machine learning model via artificial intelligence to manage medical issues was discussed. However, in recent time, the technological advances have made fundamental advancement. AI allows recognition of challenging relationships and manage large data sets, which is tasking, difficult and time consuming for humans. Furthermore, the identification and resultant risk classification that is needed for proactive surveillance trials contributes to enormous subjectivity. The use of AI models reduces the level of subjectivity, making it possible to use less reference time and improving the general efficiency and accuracy in trials. The economic encumbrance of the management which is present in prostate cancer must be acknowledged. Moreover, AI makes it feasible to accomplish this, while even improving current outcomes in active surveillance trials.

REFERENCES

Alitto, A. R., Gatta, R. and Vanneste, B. (2017). PRODIGE: Prediction models in prostate cancer for personalized medicine challenge. *Future Oncol*., 13: 2171–2181. doi:10.2217/fon-2017-0142

Antonelli, M., Johnston, E. W. and Dikaios, N. (2019). Machine learning classifiers can predict Gleason pattern 4 prostate cancer with greater accuracy than experienced radiologists. *Eur Radiol*., 29: 4754–4764. doi:10.1007/s00330-019-06244-2

Auffenberg, G. B. Ghani, K. R. and Ramani, S. (2019). Urological Surgery Improvement, 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.*, 75: 901–907. doi:10.1016/j.eururo.2018.09.050

Bhele, S., Ma, Z. and Mohanty, S. (2014). A machine learning tool to complement Gleason grading of prostate carcinoma. 2014.

Carroll, P. H. and Mohler, J. L. (2018). NCCN guidelines updates: prostate cancer and prostate cancer early detection. *J. Natl. Compr. Canc. Netw.*, 16: 620–623. doi:10.6004/jnccn.2018.0036

Choyke, P. L. (2017). Loeb S. Active surveillance of prostate cancer. *Oncology (Williston Park)*, 31: 67–70.

Fehr, D., Veeraraghavan, H. and Wibmer, A. (2015). Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proc. Natl. Acad. Sci. U S A.*, 112: E6265–6273. doi:10.1073/pnas.1505935112

Filella, X. and Foj, L. (2016). Prostate cancer detection and prognosis from Prostate Specific Antigen (PSA) to exosomal biomarkers. *Int J. Mol. Sci.*, vol. 17.

Fruge, A. D. Smith, K. S. Bail, J. R. Rais-Bahrami, S. and Demark-Wahnefried, W. (2020). Biomarkers associated with tumor Ki67 and Cathepsin L gene expression in prostate cancer patients participating in a presurgical weight loss trial. *Front Oncol.* 10: 544201. doi: 10.3389/fonc.2020.544201

Ge, P., Gao, F. and Chen, G. (2015). Predictive models for prostate cancer based on logistic regression and artificial neural network. *2015 IEEE International Conference on Mechatronics and Automation (ICMA)*. Pg. 1472–1477.

Green, W. J., Ball, G. and Hulman, G. (2016). KI67 and DLX2 predict increased risk of metastasis formation in prostate cancer – a targeted molecular approach. *Br. J. Cancer.*, 115: 236–242. doi:10.1038/bjc.2016.169

Gustavsen, G., Gullet, L., Cole, D., Lewine, N. and Bishoff, J. T. (2010). Economic burden of illness associated with localized prostate cancer in the United States. *Future Oncol.* Vol. 16, pg. 4265–4277. doi:10.2217/fon-2019-0639

Harmon, S. A. Tuncer, S., Sanford, T., Choyke, P. L. and Türkbey, B. (2019). Artificial intelligence at the intersection of pathology and radiology in prostate cancer. *Diagn Interv Radiol.*, 25: 183–188.

Humphrey, P. A. (2017). Histopathology of prostate cancer. *Cold Spring Harb Perspect. Med.*, 7(10): a030411. doi:10.1101/cshperspect.a030411

Jakobsen, N. A., Hamdy, F. C. and Bryant, R. J. (2016). Novel biomarkers for the detection of prostate cancer. *J. Clin. Urol.*, 9: 3–10. doi:10.1177/2051415816656121

Jin, W., Fei, X., Wang, X., Song, Y. and Chen, F. (2020). Detection and prognosis of prostate cancer using blood-based biomarkers. *Mediators Inflamm.* 2020: 8730608. doi: 10.1155/2020/8730608

Kim, Y., Jeon, J. and Mejia, S. (2016). Targeted proteomics identifies liquid-biopsy signatures for extracapsular prostate cancer. *Nat. Commun*., 7: 11906.

Kumar, N., Verma, R., Arora, A., Kumar, A., Gupta, S., Sethi, A. and Gann, H. P. (2017). Convolutional neural networks for prostate cancer recurrence prediction. Proceedings of Medical Imaging: Digital Pathology. SPIE Medical Imaging, Orlando, Florida, United States. vol. 10140 https://doi.org/10.1117/12.2255774

Lawrentschuk, N., Lockwood, G. and Davies, P. (2011). Predicting prostate biopsy outcome: artificial neural networks and polychotomous regression are equivalent models. *Int. Urol. Nephrol*., 43: 23–30. doi:10.1007/s11255-010-9750-7

Martignano, F., Rossi, L. and Maugeri, A. (2017). A-based biomarkers for prostate cancer detection. *Clin. Chim. Acta*., 473: 96–105. doi:10.1016/j.cca.2017.08.009

McDonald, A. C., Vira, M. and Shen, J. (2018). Circulating microRNAs in plasma as potential biomarkers for the early detection of prostate cancer. *Prostate*, 78: 411–418.

Moore, C. M., Petrides, N. and Emberton, M. (2014). Can MRI replace serial biopsies in men on active surveillance for prostate cancer? *Curr Opin Urol*., 24: 280–287. doi:10.1097/MOU.0000000000000040

Neeraj, K., Ruchika, V. and Ashish, A. (2017). Convolutional neural networks for prostate cancer recurrence prediction. *Proc. SPIE,*

Park, S. and Byun, J. (2021). A study of predictive models for early outcomes of postprostatectomy incontinence: Machine learning approach vs. logistic regression analysis approach. *Appl. Sci*., 11(13): 6225.

Strom, P., Kartasalo, K. and Olsson, H. (2020). Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based. *Diagnostic Study, Lancet Oncol*., 21: 222–232. doi: 10.1016/S1470-2045(19)30738-7

Stojadinovic, M., Stojadinovic, M. and Pantic, D. (2018). Decision tree analysis for prostate cancer prediction. *Srp Arh Celok Lek*., 147: 39.

Torre. L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J. and Jemal, A. (2012). Global cancer statistics, 2012. *C A Cancer J. Clin*., 65: 87–108. doi:10.3322/caac.21262

Toivonen, J., Montoya Perez, I. and Movahedi, P. (2019). Radiomics and machine learning of multisequence multiparametric prostate MRI: towards improved non-invasive prostate cancer characterization. *PLoS One*, 14: e0217702. doi: 10.1371/journal.pone.0217702

Waliszewski, P., Wagenlehner, F., Gattenlöhner, S. and Weidner, W. (2015). On the relationship between tumour structure and complexity of the spatial distribution of cancer cell nuclei: a fractal geometrical model of prostate carcinoma. *Prostate*, 75: pg. 399–414. doi: 10.1002/pros.22926