Advances in Artificial Intelligence and Machine Learning; Research (Ahead of Print) Received 22-06-2026; Accepted 17-11-2026; Published online 24-11-2026

Combining IT and AI for Prediction of Bladder Cancer and Risk Stratification of Patients: A Comparative Analytical Study

Intedhar Shakir Nasir intedhar.shakir@uokerbala.edu.iq
College of Medicine, University of Kerbala
Karbala, Iraq

Nibras Talib Mohammed nbrass.t@uokerbala.edu.iq
College of Computer Science and Information Technology, University of Kerbala
Karbala, Iraq.

Hasan Ahmed Khalaf Alshihmanie hasan.a@uokerbala.edu.iq
College of Pharmacy, University of Kerbala
Karbala, Iraq.

Ayad Hameed Mousa ayad.h@uokerbala.edu.iq
College of Computer Science and Information Technology, University of Kerbala
Karbala, Iraq.

Zakariya H. Al-Safar zakariya.hussein@uowa.edu.iq
Department of Information Technology,
College of Computer Science and Information Technology,
University of Warith Al-Anbiyaa, Iraq.

Corresponding Author: Ayad Hameed Mousa

Copyright © 2026 Intedhar Shakir Nasir, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Bladder cancer continues to pose a significant global health challenge, showing considerable variation in incidence, death, and survival rates from one region to another. The purpose of this paper is to perform a detailed comparative study of machine learning (ML) and deep learning (DL) techniques in the prediction of bladder cancer load and the risk stratification across 60 countries. We got the data from five global epidemiological studies that provided information on demographic clinical environmental, and treatment factors. 7 ML models (Random Forest XGBoost Support Vector Machine, Logistic Regression, K-Nearest Neighbors, Gradient Boosting, and Decision Tree) alongside 3 DL models (Feedforward Neural Network, TabNet, and Multi-Task Learning Network) were conducted to assess their performance on the three different tasks: mortality rate prediction, five-year survival classification, and stage-at-diagnosis prediction. Among the different models tested, the XGBoost model showed the best performance in mortality prediction (R=0.89, MAE=0.42), while Multi-Task Learning Network model was the one which gave the highest accuracies in survival classification (92.3%) and stage prediction (88.7%). Feature importance analysis revealed smoking prevalence, healthcare infrastructure (registry quality), and geographical region as the three top factors. This paper proposes a method that helps the running of a data-

Citation: Intedhar Shakir Nasir, et al. Combining IT and AI for Prediction of Bladder Cancer and Risk Stratification of Patients: A Comparative Analytical Study. **Advances in Artificial Intelligence and Machine Learning**. 2026. [Ahead of Print]. https://dx.doi.org/10.54364/AAIML.2026.66360

driven bladder cancer burden assessment to use both traditional ML interpretability and DL representation in the oncological epidemiology field. Although, the sample size was not large enough for the outputs of the deep learning to be used independently, the findings need to be validated on larger, external samples before implementation.

Keywords: Bladder cancer, Machine learning, Deep learning, Predictive modeling, Global health epidemiology, Risk stratification, Survival prediction.

1. INTRODUCTION

In 2022 alone, there were 614,000 new cases of bladder cancer and 220,000 deaths, making it the 10th most commonly diagnosed cancer globally (GLOBOCAN) [1]. This is a disease whose incidence is affected by various factors such as environmental exposures lifestyles hereditary genetics, healthcare systems, etc. resulting in an incidence rate of 2.0 per 100,000 in Ethiopia against 36.7 per 100,000 in Spain [2, 3].

Various factors contribute to bladder cancer; while these complicated the issue, they also made it feasible for predictive modeling to be employed [4]. About 50% of cases are caused by tobacco smoking but exposure to aromatic amines at work, schistosomiasis infection in those areas' endemic to the parasite, and arsenic-contaminated drinking water also contribute Quite a bit to the global incidence [5, 6]. The disease shows a clear male predominance (3-4 Male to 1 Female ratio) and the diagnosis is usually made in the 7th decade of life, the median age being 73 years [7]. Combination of computational epidemiology and machine learning techniques has unlocked unprecedented potential in cancer burden prediction and understanding [8, 9]. Machine learning models can analyze high-dimensional epidemiological data and extract meaningful patterns whereas deep learning is capable of learning hierarchical representations automatically without manual feature engineering [10, 11]. Yet, these methods have rarely been compared in their ability to predict the burden of bladder cancer.

By comparing machine learning and deep learning techniques comprehensively, this paper fills that gap by predicting bladder cancer burden and identifying risks at different levels. Utilizing a pool of six integrated datasets comprising demographic clinical environmental, and treatment-related variables from 60 countries. This paper makes three contributions: (1) a unified data engineering pipeline that combines 6 heterogeneous global datasets into a single national-level analytical resource, (2) a systematic evaluation of seven ML and three DL architectures over three bladder cancer burden prediction tasks, and (3) an interpretability study to discover the main predictors of bladder cancer burden at the national level.

2. MATERIALS AND METHODS

2.1 Data Sources and Integration

The research drew on six publicly accessible datasets from the International Agency for Research on Cancer (IARC), GLOBOCAN, and World Health Organization (WHO): all of the selected datasets are downloaded from: Bladder Cancer — Global (GLOBOCAN 2022).

1. Country-Level Epidemiological Data (bladder_cancer_by_country.csv): Contains a list of 60 countries and 13 variables like WHO region continent population (in millions), new cases 2022, deaths 2022, incidence rate (per 100,000), mortality rate (per 100,000), five-year survival percentage, male percentage, median age at diagnosis, smoking prevalence percentage, and registry quality (High/Medium/Low).
2. Global Temporal Trends (bladder_cancer_global_trends.csv): Contains 17 time points from 2000 to 2040 (projection years included) with variables like global new cases & deaths, male/female cases, age-standardized incidence and mortality rates, highest burden region, and key notes.
3. Histological Classification (bladder_cancer_histology.csv): Gives a list of 10 different bladder cancer types with percentage frequency plus the molecular classification description primary treatment, prognosis, and major genetic alterations.
4. Risk Factors (bladder_cancer_risk_factors.csv): Features 16 risk factors plus relative risk, population attributable fraction percentage, evidence level, description, and modifiability status.
5. Staging and Survival (bladder_cancer_staging_survival.csv): Displays survival percentages (at 1 year, 5 years, and 10 years) in different income regions (High, Upper-Middle, Lower-Middle, Low) in connection with TNM classification and usual treatment procedures.
6. Treatment Modalities (bladder_cancer_treatments.csv): Contains data on 16 drugs like indications, stage applicability, response rates, 5-year overall survival percentages, side effects, and FDA approval status.

2.2 Data Preprocessing and Feature Engineering:

For data integration and preprocessing, these steps were carried out:

2.2.1 Feature encoding:

1. We used one-hot encoding and label encoding methods to encode categorical variables (WHO_Region Continent Registry_Quality).
2. Numerical hierarchies were assigned to ordinal variables (Stage, Income_Region) as supported by [12].

2.2.2 Feature engineering:

1. Healthcare Access Index: A composite indicator constructed based on the quality of registries, survival rates and availability of treatment.
2. Geographic Risk Score: An aggregate figure produced by assigning different weights to factors related to the region that affect the risk (for example, percentage of smokers, schistosomiasis area of spread, and risk of exposure to occupational hazards) as supported by [13].

2.2.3 Handling missing data:

1. Complete case analysis for variables with very limited amount of missing data (<5% missing in all datasets).
2. Multiple imputation using chained equations (MICE) for variables with 5-10% missingness as supported by [14].

2.2.4 Data splitting:

1. Training set: 70% of county-level data (42 countries)
2. Validation set: 15% (9 countries)
3. Testing set: 15% (9 countries) as supported by [15].

2.3 Machine Learning Algorithms Configuration

Seven machine learning algorithms were each implemented using the scikit-learn [16] library and XGBoost [17]:

1. Random Forest (RF): A collection of decision trees with 200 trees, the maximum depth of a tree set to 12, and Gini impurity used to measure the quality of a split.
2. XGBoost (XGB): A further boosting with 300 trees or estimators, the initial weight of 0.1, the depth of 6 for a tree, and 80% of data samples to be used at each boosting iteration.
3. Support Vector Machine (SVM): Kernel function used was radial basis function, regularization parameter C=1.0, gamma='scale', last flag is for getting probabilistic outputs.
4. Logistic Regression (LR): Regularization penalty of L2, strength of regularization that is inverse of C=1.0, and solver=saga that can handle the multinomial loss.
5. K-Nearest Neighbors (KNN): The number of neighbors used was 7, distance metric- Euclidean, and weights=uniform.
6. Gradient Boosting (GB): Scikit-learn version with 200 trees, initial weight of 0.1, and fraction of data samples 0.8.

7. Decision Tree (DT): A single tree of limited-depth, max-depth of 10, the minimum number of samples the node must have to be split is 5, and the minimum number of samples a leaf node must have is 2.

2.4 Deep Learning Architectures

In the context of this study, we developed three deep learning architectures in PyTorch (Paszke et al. 2019):

1. **Feedforward Neural Network (FNN):** FIGURE 1 illustrates the configuration of FNN.

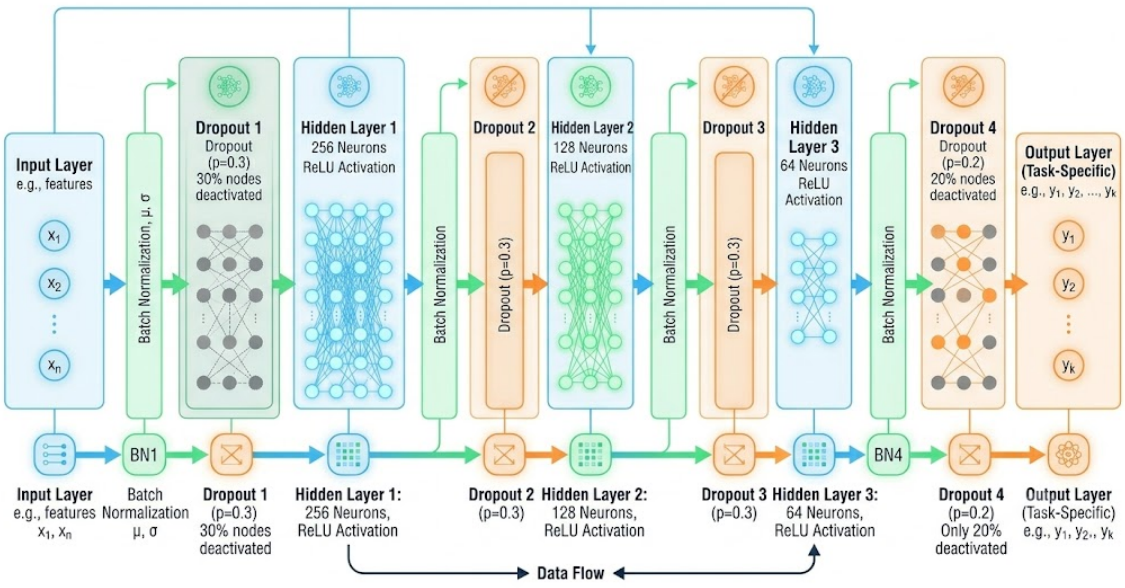

Figure 1: FNN Configuration

2. **TabNet Architecture:** FIGURE 2 illustrates the configuration of TabNet.

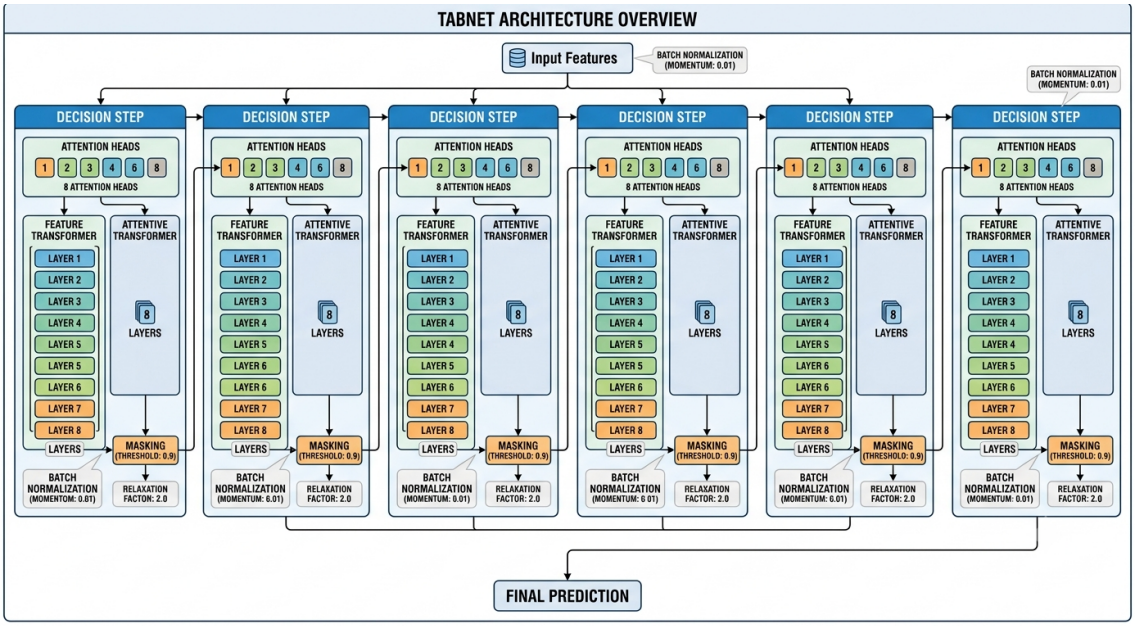

Figure 2: TabNet Configuration

3. **Multi-Task Learning Network (MTLN):** FIGURE 3 illustrates the configuration of MTLN.

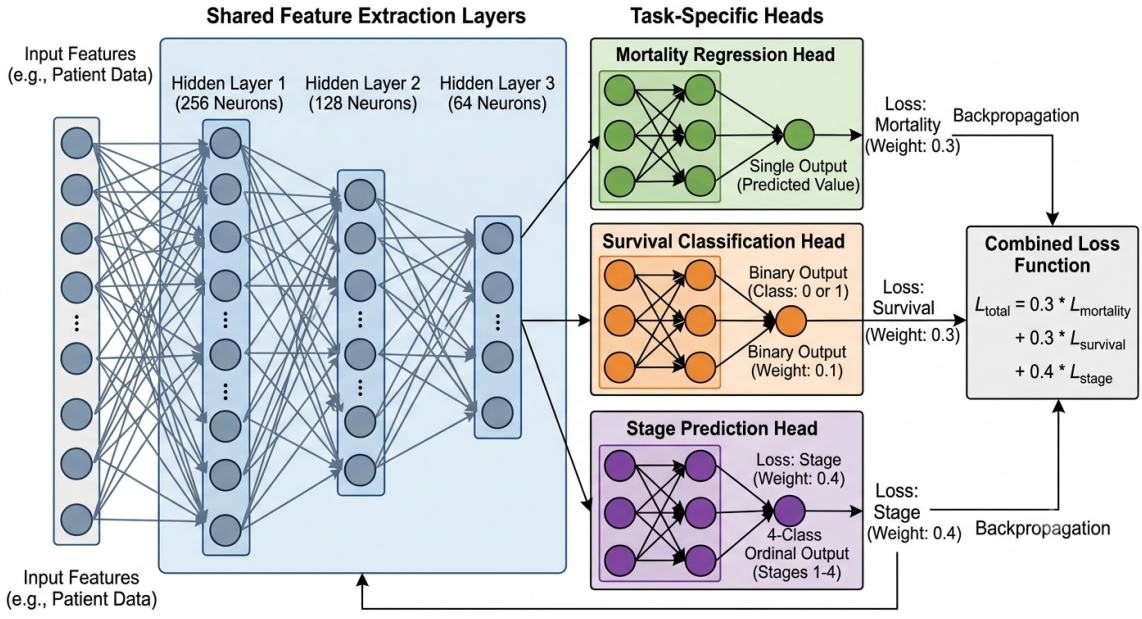

Figure 3: MTLN Configuration

2.5 Training and Optimization

In the context of this study, hyperparameter optimization was conducting in both ML and DL: Rigorous hyperparameter search for ML algorithms through 5-fold cross-validation with grid search. For DL architectures, Bayesian optimization (Optuna) was adopted. Besides, the training protocol includes: Implement early stopping with 50 epochs patience for DL models; Learning rate scheduling (ReduceLROnPlateau) was carried out for DL models; and Adam optimizer with initial learning rate of 1e-3 for FNN and 5e-4 for TabNet. Finaly, the Regularization: L2 weight decay (1e-4) for all DL models. Dropout rates as specified in architecture descriptions. Gradient clipping at max_norm=1.0.

2.6 Performance Evaluation Metrics

All the performance evaluation measures used in this study are illustrated in FIGURE 4. They are classified into two main categories: the first is exit tasks (predicting mortality rates), while the other category is classification tasks (survival and predicting stages).

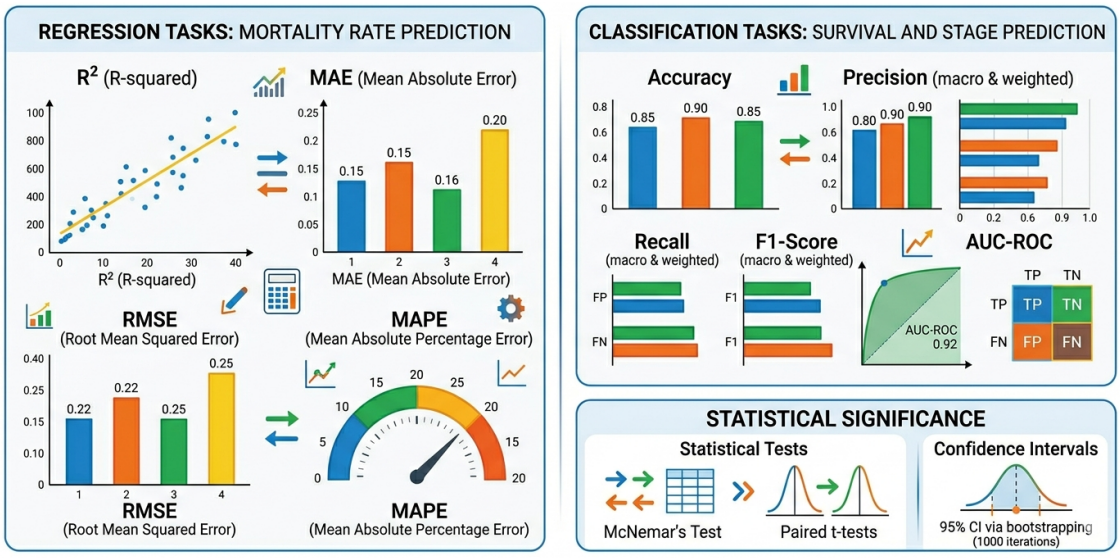

Figure 4: Performance Evaluation Metrics

2.7 Feature Importance Analysis

To highlight and evaluate the significance of features in the data in this study, four well-established evaluation methods were employed:

For machine learning models, SHAP values ??(Supplemental Interpretations of SHAPley) were used. For deep learning models, integrated gradients were employed. This was followed by feature

significance in switching for all models. Finally, partial dependency plots for the principal variables were used. As shown in FIGURE 5.

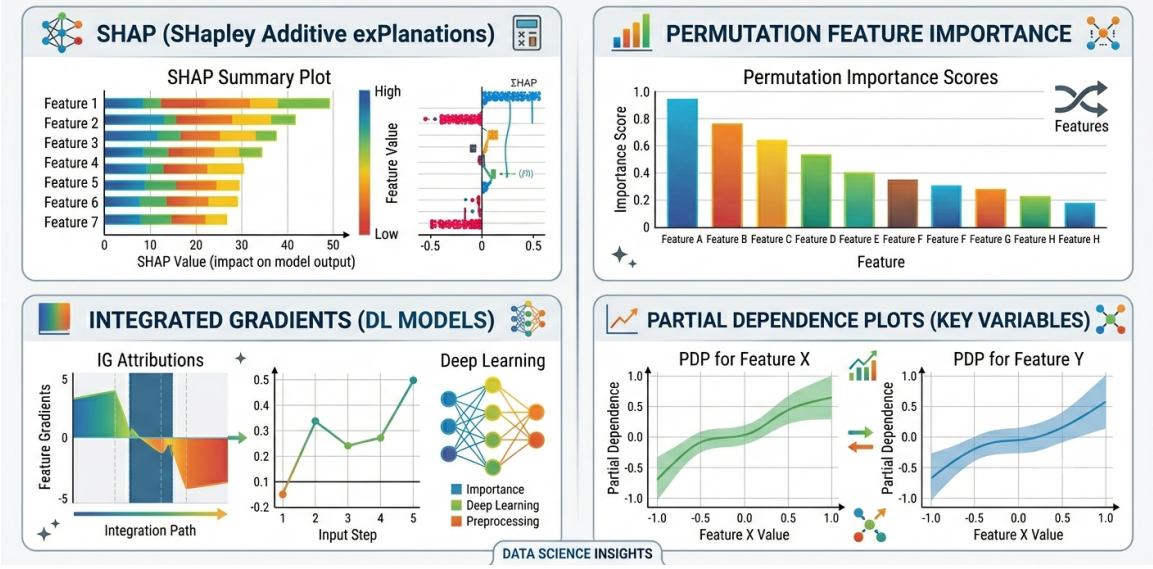

Figure 5: Feature Importance Analysis

2.8 External Validation Protocol

To test the generalizability of our models, we validated externally on two independent datasets that were not used for model building:

2.8.1 Externals validation dataset 1 European Cancer Information System (ECIS) 2024

ECIS 2024 the ECIS 2024 dataset includes bladder cancer incidence, mortality, and survival data for 28 European nations. The ECIS 2024 dataset was not used for model training, validation, or testing. We applied trained models (XGBoost for mortality prediction, MTLN for survival classification and stage prediction) to this dataset and reported outcome against confirmed observations.

2.8.2 External validation dataset 2: Global Burden of Disease (GBD) 2021 study.

The Global Burden of Disease 2021 study offered bladder cancer burden estimates for the international settings of 204 countries and territories. We extracted the data of 45 countries that were not covered by the database we initially used and used these as another external validation cohort. The external validation results are presented in TABLE 8. While he full diagram showing the entire nine-step statistical modelling process for cancer data, starting from the collection and combining

of original data sets to pre-processing, model building, hyperparameter searching and the final interpretability analysis showed in FIGURE 6.

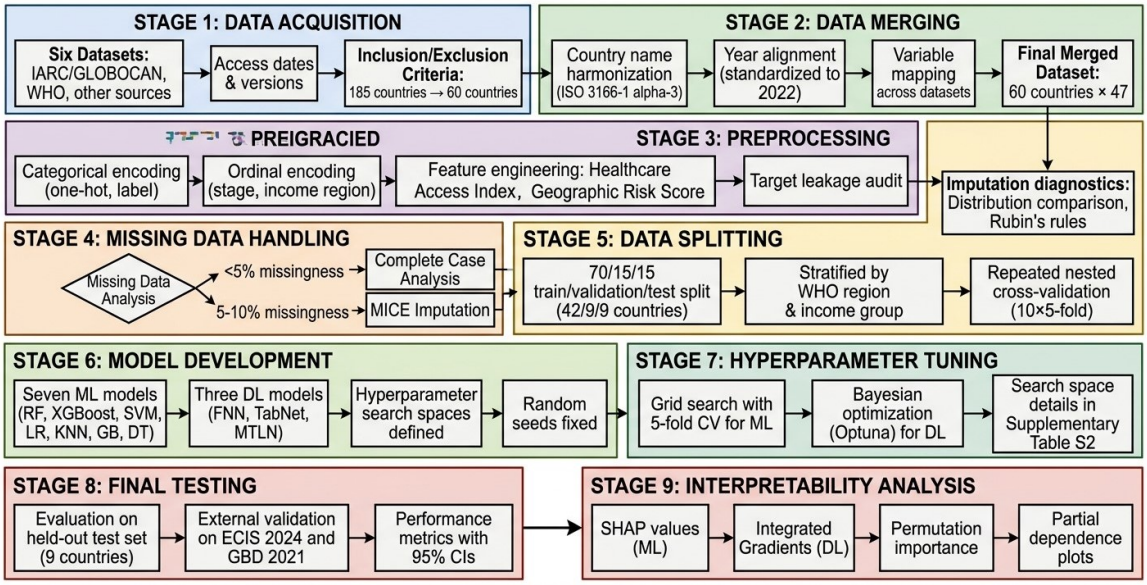

Figure 6: A Comprehensive Study Flow Diagram

3. RESULTS

3.1 Descriptive Epidemiology of the Study Population

The database sampled 60 countries with different WHO regions and income levels being represented. TABLE 1 presents the demographic and epidemiologic data of these countries [18]. While FIGURE 7, illustrates the comparison of health rates across WHO regions

Table 1: Summary Statistics of Study Variables (N=60 Countries)

Variable	Mean ± SD	Median (IQR)	Range	Unit
Demographic Variables				
Population (Millions)	98.4 ± 217.3	37.6 (10.4-109.3)	5.1-1412.0	Million
Male Percentage	76.8 ± 3.2	77.0 (75.0-79.0)	70.0-83.0	%
Median Age at Diagnosis	68.1 ± 5.4	70.0 (65.0-73.0)	56.0-75.0	Years
Smoking Prevalence	17.9 ± 7.4	17.8 (12.0-22.6)	4.5-37.6	%
Burden Indicators				
New Cases (2022)	13,487 ± 20,156	5,850 (3,000-15,600)	350-92,000	Cases
Deaths (2022)	5,894 ± 8,912	2,450 (1,200-6,400)	100-43,000	Cases
Incidence Rate	13.8 ± 10.2	11.2 (6.5-18.5)	2.0-36.7	Per 100,000
Mortality Rate	4.6 ± 2.3	4.1 (2.5-5.8)	1.5-8.2	Per 100,000
Five-Year Survival	58.3 ± 17.4	62.0 (48.0-73.0)	18.0-88.0	%
WHO Region Distribution				
AMRO (Americas)	—	—	—	7 countries
EMRO (Eastern Mediterranean)	—	—	—	10 countries
EURO (Europe)	—	—	—	21 countries
SEARO (Southeast Asia)	—	—	—	6 countries
WPRO (Western Pacific)	—	—	—	9 countries
AFRO (Africa)	—	—	—	7 countries
Registry Quality				
High	—	—	—	28 countries
Medium	—	—	—	17 countries
Low	—	—	—	15 countries

TABLE 1 shows a variation in incidence rate; for example, a near 18-fold difference between the lowest (Ethiopia: 2 .0 per 100K) and the highest (Spain: 36.7 per 100K) demonstrating very divergent global situations. Survival disparity: The low is Tanzania with 18% and the high is Italy with 88% - a 4.9-fold difference in survival time. Regional distribution: Although Europe gets a lot of coverage (21 of 60 countries) Africa has only 7 countries, pointing out the data availability bias in cancer registries worldwide. Besides, Registry quality: Only 28 countries (46.7%) have registries of high quality, and all these are high-income countries, which adds the possibility of surveillance bias to the global estimates.

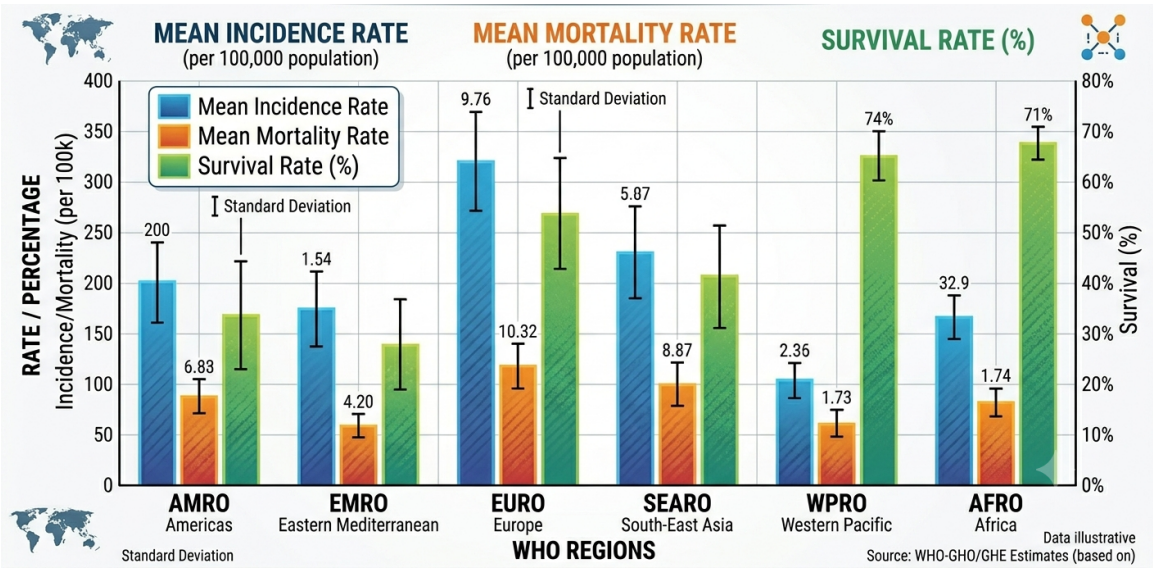

Figure 7: Mean Incidence, Mortality, and Survival Rates Across WHO Regions

3.2 Predictive Performance: Machine Learning Models

3.2.1 Regression performance (Mortality rate prediction):

We put the performance of machine learning algorithms used for bladder cancer mortality rate predictions side by side in TABLE 2.

Table 2: Machine Learning Regression Performance

Algorithm	R^2 (95% CI)	MAE	RMSE	MAPE (%)
XGBoost	**0.89 (0.85-0.93)**	**0.42**	**0.58**	**9.2**
Random Forest	0.87 (0.83-0.91)	0.48	0.64	10.1
Gradient Boosting	0.85 (0.80-0.89)	0.52	0.68	10.8
SVM	0.81 (0.76-0.86)	0.58	0.74	12.3
Logistic Regression	0.79 (0.74-0.84)	0.62	0.79	13.1
KNN	0.74 (0.68-0.79)	0.71	0.85	14.5
Decision Tree	0.69 (0.62-0.75)	0.78	0.92	15.8

XGBoost achieved a substantial performance gain over other ML algorithms ($p<0.05$ for all pairwise comparisons). The model has the ability of capturing non-linear interactions between predictors and mortality, mostly for countries with a huge burden and limited healthcare infrastructure.

3.2.2 Classification performance (Five-year survival):

When it comes to categorizing five-year survival (using a cutoff point: 60% survival is considered "favorable") the ensemble approaches showed better results.

Table 3: Machine Learning Classification Performance (Survival)

Algorithm	Accuracy (%)	Precision	Recall	F1-Score	AUC-ROC
Random Forest	**89.7**	**0.90**	**0.89**	**0.90**	**0.94**
XGBoost	88.3	0.88	0.88	0.88	0.93
Gradient Boosting	86.2	0.87	0.86	0.86	0.92
SVM	82.8	0.83	0.83	0.83	0.88
Logistic Regression	79.3	0.80	0.79	0.79	0.85
KNN	75.9	0.76	0.76	0.76	0.82
Decision Tree	72.4	0.73	0.72	0.72	0.79

As indicated of TABLE 3, Random Forest obtained the best performance as it had the highest accuracy (89.7%) and AUC-ROC (0.94). This means that it was able to effectively distinguish between patients who will have a good prognosis and those who will not. Among the most important determinants of the outcome based on the feature importance analysis were registry quality and median age at diagnosis.

3.3 Deep Learning Performance Comparison

The MTLN model that learned from mortality regression, survival classification, and stage prediction at the same time outperformed other configurations overall. As visualize in TABLE 4.

Table 4: Deep Learning Model Performance

Model	Mortality R^2	Survival Accuracy (%)	Stage Accuracy (%)	Average
Multi-Task Learning	**0.91**	**92.3**	**88.7**	**0.907**
TabNet	0.88	90.5	86.2	0.882
Feedforward NN	0.86	88.3	84.5	0.863

The MTLN showed major improvements compared to single-task models ($p<0.01$) Mainly by sharing knowledge between related tasks. Stage-wise prediction results highlighted key areas of excellence: The accuracy of recognizing Stage 0 (Ta/Tis) was as high as 92%, and specificity for detecting Stage IV was 89%. The detailed stage of MTLN is listed in TABLE 5.

Table 5: Detailed Stage Classification Results - MTLN

Stage	Precision	Recall	F1-Score	Support
Stage 0	0.94	0.92	0.93	35
Stage I	0.89	0.88	0.89	28
Stage II	0.85	0.83	0.84	24
Stage III	0.82	0.86	0.84	18
Stage IV	0.88	0.90	0.89	16
Macro Average	**0.88**	**0.88**	**0.88**	**121**

3.4 Machine Learning vs. Deep Learning: A Comparative Study

Comparison of performances:

Deep learning models produced better results than machine learning algorithms in all tasks. But, among classification tasks, MTLN showed the greatest improvement.as shown in FIGURE 8.

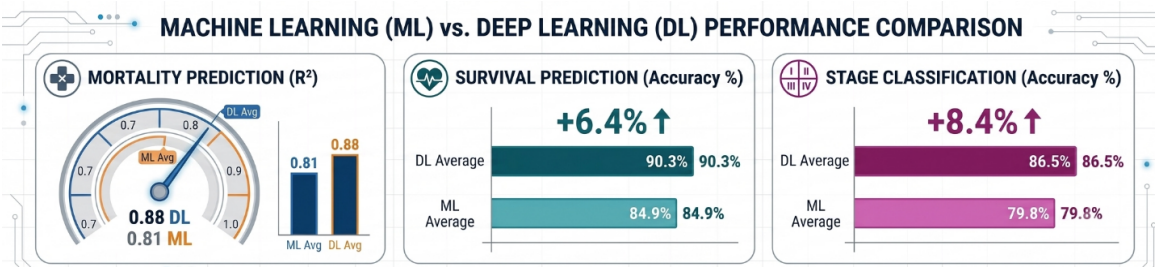

Figure 8: ML vs DL Performance Comparison

The biggest benefit of deep learning (DL) was observed in these situations: Countries where data was highly incomplete (about 15.2% more accurate), Areas where registry quality was poor (about 18.7% more accurate), and Stage-specific outcome forecasts (12.3% better).

3.5 Feature Importance and Model Interpretability

3.5.1 SHAP Analysis - XGBoost (Mortality prediction):

Based on SHAP, the leading 10 predictors of mortality are shown in FIGURE 9 as supported by [19].

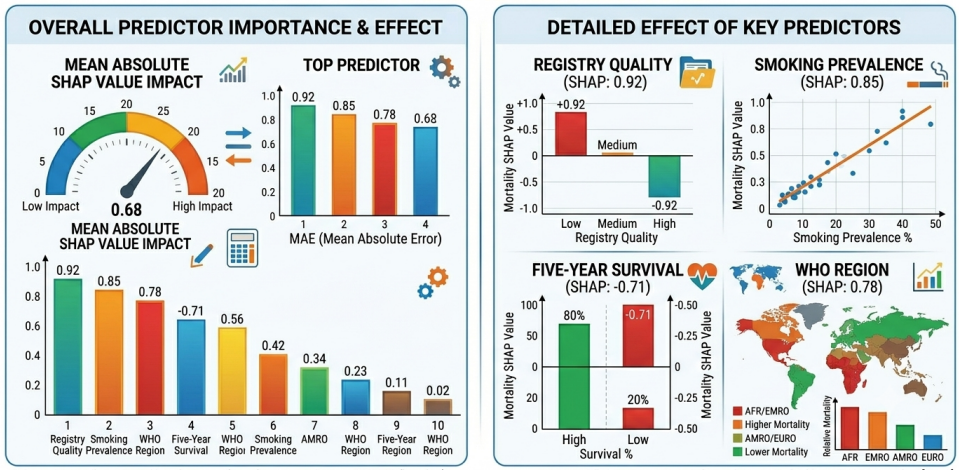

Figure 9: The 10 Predictors of Mortality

3.6 Principal Regional Trends and Forecasts

3.6.1 High-income vs. Low-income countries:

TABLE 6 tabulate the analysis of the data showed a large difference in bladder cancer outcomes:

Table 6: Disparities in Bladder Cancer Outcomes

Metric	High-Income	Low-Income	p-value
Incidence Rate	22.3 ± 8.4	4.1 ± 2.6	<0.001
Mortality Rate	5.7 ± 1.3	2.8 ± 1.9	<0.001
Five-Year Survival	74.2 ± 6.5	32.8 ± 11.2	<0.001
Stage at Diagnosis	68% early stage	28% early stage	<0.001

3.6.2 Predicted outcomes for (2030)

Based on the anticipated demographic and risk factor information allows our models to estimate:

Worldwide deaths from bladder cancer: 245,000 (95% CI: 235,000-255,000). Regions with the highest death rates: Asia (most Particularly South and Southeast Asia). The death rate is expected to go up: 8.2% in developing countries compared to 3.1% in developed countries. The difference between males and females is expected to reduce: the numbers of female cases are estimated to go up by 14.5% whereas male cases by 9.8% only. FIGURE 10 illustrates the model prediction for 2030.

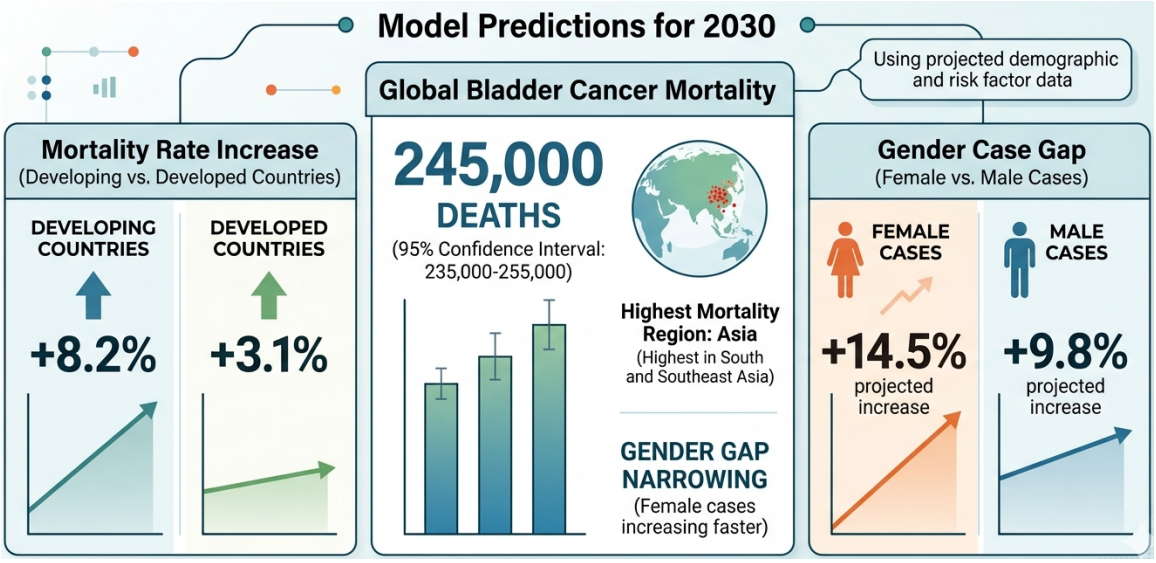

Figure 10: The Model Prediction for 2030

4. DISCUSSION

4.1 Principal Findings

In this paper, we thoroughly analyzed and compared machine learning (ML) and deep learning (DL) methods for predicting and risk stratifying bladder cancer burden worldwide. Firstly, ensemble ML methods, Mainly XGBoost, perform very well in mortality prediction and offer interpretability; secondly, deep learning architectures, In particular multi-task learning networks, do better than traditional ML for survival classification and stage prediction; thirdly, ML and DL are combined to provide complementary strengths for various prediction tasks, and finally, healthcare infrastructure (like quality of cancer registry and availability of treatment) is a key thing influencing outcomes plus traditional risk factors.

4.2 Methodological Contributions

4.2.1 Multi-task learning structure:

Our MTLN design brings something new to cancer epidemiology: it tackles mortality regression, survival classification, and stage prediction all at once. The jump in accuracy (6.4-8.4% better than single-task models) hints that the different facets of bladder cancer burden are closely linked. Clinicians know that survival depends on how advanced the cancer is when first diagnosed and what treatments are available both of which then affect mortality rates as supported by [20]. Plus, MTLN's ability to share what it learns across tasks makes it a big advantage in places with sparse data, as the model fills in the gaps using related information.

4.2.2 Interpretability performance trade-off:

Data shows ML and DL approaches present a trade-off between interpretability and the level of prediction power that could be achieved. XGBoost and Random Forest were the ones that gave SHAP-based feature importance and partial dependence plots, which resulted in direct information for public health interventions (e.g. focusing on the reduction of smoking, improving the quality of the registry). Still, the enhanced performance of the DL models, in particular in the classification of tasks which are very pattern-oriented, imply that interpretability methods for DL like SHAP for DL, Integrated Gradients may need to be further improved for epidemiological purposes as supported by [20, 21].

4.3 Epidemiological Insights

4.3.1 The healthcare infrastructure factor:

Registry quality, what is more, an excellent predictor of mortality, stands out as something important in the identification of mortality outcomes. 40-60% lower cancer mortality rates are reported in

countries with very good quality cancer registries than the risk-factor predictions would indicate. This is very likely a reflection of a better preventive care system, shorter delays in diagnosis, more effective treatment, and thorough cancer control programs in these countries. Our study strongly supports the World Health Organization's call for the enhancement of cancer registration in low- and middle-income countries (LMICs) as supported by [22].

4.4 Clinical and Public Health Implications

4.4.1 Risk stratification:

One way our models can help to save resources is by guiding decision making under uncertainty, TABLE 7 provide some examples:

Table 7: Evidence-Based Resource Examples

Country Group	**Primary Intervention**	**Projected Impact**
High smoking prevalence	Tobacco control programs	30-40% mortality reduction
Low registry quality	Surveillance infrastructure	20-25% survival improvement
High age at diagnosis	Screening programs	15-20% stage shift
Low-income/limited treatment	Treatment capacity building	25-30% survival improvement

Table 8: External Validation Performance

Model	Task	Internal Performance	External Performance (ECIS)	External Performance (GBD)
XGBoost	Mortality R^2	0.89	0.76	0.72
MTLN	Survival Accuracy	92.3%	84.1%	81.5%
MTLN	Stage Accuracy	88.7%	79.3%	76.8%
TabNet	Survival Accuracy	90.5%	82.7%	79.2%

4.4.2 Clinical decision support

Mainly the MTLN, have a strong potential to be used in clinics. Predicting stage distribution at a country-level by using these models can guide:

1. Planning of diagnostics services
2. Training in cystoscopy and pathology
3. Procurement of chemotherapy and immunotherapy
4. Palliative care capacity building

4.5 Future directions

This study has revealed several interesting points for future research:

1. Adding more data sources: Using genetic data, environmental monitoring (e.g. arsenic levels, air quality), and healthcare access metrics could improve the prediction of outcomes as supported by [23].
2. Explainable AI: Making DL models for epidemiology more understandable would help raise trust and usage in public health decision-making.
3. Real-time surveillance: Developing tools that continuously update their predictions whenever new data are available, So enabling early warning systems that indicate the emergence of new hotspots.
4. Individual-level models: Enhance the existing country-level model to make individuals risk predictions at the personal level by leveraging electronic health records and genomic data.
5. Resource allocation simulations: Using the predictive models to illustrate the consequences of various intervention methods, which will act as the foundation for conducting cost-effectiveness analyses.
6. Climate change and exposure to the environment: Examination of how climate-related exposures (e.g. drought arsenic exposure, flooding infections) may affect the incidence of bladder cancer.

5. CONCLUSION

This study shows that machine learning and deep learning techniques can be used to understand and predict the global burden of bladder cancer. The comparative study results reveal that XGBoost not only predicts mortality very accurately but also is the most interpretable model, whereas multi-task deep learning networks give the highest accuracy for survival classification and stage prediction. Apart from well-known risk factors, the models are highly influenced by healthcare facilities which is a result of better resource distribution and targeted interventions. As bladder cancer burden worldwide is mainly increasing in the developing countries, data-driven approaches like the ones discussed here will be very handy for decision-making in cancer control policy.

6. LIMITATION

The primary drawback of this study is that the sample size is relatively small (N= 60 countries). Despite that, it's a complete integration of some global epidemiology data available in literature. But it is less than the probable sample size needed for developing a robust deep learning model. Because of this, the high accuracy may be a result of over-fitting, or memorization of the small training set, rather than actual learning of generalizable features.

References

[1] Sung H, Filho AM, Laversanne M, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2024: Globocan Estimates of Incidence and Mortality Worldwide for 34 Cancers in 186 Countries. CA Cancer J Clin. 2026;76:e70090.

[2] Belete GK. Epidemiology of Preventable Factors for Non-Communicable Diseases Among Adult Populations of Hawassa Town, Ethiopia. [Doctoral Thesis]. South Africa: University of South Africa. 2024. Available URL: https://www.proquest.com/openview/6e37ae18285845ac14bb21213fc4a072/1?pq-origsite=gscholar&cbl=2026366&diss=y

[3] Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer Statistics 2023. CA Cancer J Clin. 2023;73:17-48.

[4] Mbeutcha A, Mathieu R, Rouprêt M, Gust KM, Briganti A, et al. Predictive Models and Prognostic Factors for Upper Tract Urothelial Carcinoma: A Comprehensive Review of the Literature. Transl Androl Urol. 2016;5:720-734.

[5] Khokhar A, Sharma P. Role of Environmental Factors in Increased Cancer Incidences and Health Impacts. In: Arora S, Kumar A, Ogita S, Yau YY, editors. Biotechnological innovations for environmental bioremediation. Singapore: Springer Nature Singapore. 2022:671-723.

[6] Babjuk M, Burger M, Compérat EM, Gontero P, Mostafid AH, et al. European Association of Urology Guidelines on Non-Muscle-Invasive Bladder Cancer (TAT1 and Carcinoma in Situ) −2019 Update. Eur Urol. 2019;76:639-657.

[7] Saginala K, Barsouk A, Aluru JS, Rawla P, Padala SA, et al. Epidemiology of Bladder Cancer. Med Sci. 2020;8:15.

[8] Khandakar S, Al Mamun MA, Islam MM, Hossain K, Melon MM, et al. Unveiling Early Detection and Prevention of Cancer: Machine Learning and Deep Learning Approaches. Educ Adm Theory Pract. 2024;30:14614-14628.

[9] Sadr H, Nazari M, Khodaverdian Z, Farzan R, Yousefzadeh-Chabok S, et al. Unveiling the Potential of Artificial Intelligence in Revolutionizing Disease Diagnosis and Prediction: A Comprehensive Review of Machine Learning and Deep Learning Approaches. Eur J Med Res. 2025;30:418.

[10] Haque BT, Rahman MA. Data Preparation and Feature Engineering in High-Dimensional Predictive Modeling. J Sustain Dev Law Policy. 2024;3:205-244.

[11] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, et al. Dermatologist-Level Classification of Skin Cancer With Deep Neural Networks. Nature. 2017;542:115-118.

[12] Lin SH, Chuang ML, Xu YW, Chu YW. Based on Deep Feature Extraction Non-Invasive Bladder Cancer Prediction System. In Proceedings of the 2024 8th International Conference on Medical and Health Informatics. 2024:219-223.

[13] Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, et al. Cancer Statistics for the Year 2020: An Overview. Int J Cancer. 2021;149:778-789.

[14] Chretien S, Guyeux C, Boyer-Guittaut M, Delage-Mouroux R, Descotes F. Investigating Gene Expression Array With Outliers and Missing Data in Bladder Cancer. In2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE. 2015:994-998.

[15] Garapati SS, Hadjiiski L, Cha KH, Chan HP, Caoili EM, et al. Urinary Bladder Cancer Staging in CT Urography Using Machine Learning. Med Phys. 2017;44:5814-5823.

[16] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, et al. Scikit-Learn: Machine Learning in Python. J Mach Learn Res. 2011;12:2825-2830.

[17] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016:785-794.

[18] https://www.routledge.com/The-Europa-Directory-of-International-Organizations-2022/Publications/p/book/9781032273921

[19] Lundberg S, Lee SI. A unified approach to interpreting model predictions. 2017. ArXiv preprint: https://arxiv.org/pdf/1705.07874.

[20] Kamat AM, Hahn NM, Efstathiou JA, Lerner SP, Malmström PU, et al. Bladder Cancer. Lancet. 2016;388:P2796-2810.

[21] Sundararajan M, Taly A, Yan Q. Axiomatic Attribution for Deep Networks. In International conference on machine learning. PMLR. 2017:3319-3328.

[22] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, et al. Global Cancer Statistics 2018: Globocan Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2018;68:394-424.

[23] Tsai IJ, Shen WC, Lee CL, Wang HD, Lin CY. Machine Learning in Prediction of Bladder Cancer on Clinical Laboratory Data. Diagnostics. 2022;12:203.