



OPEN ACCESS

EDITED BY

Kyo Chul Koo,
Yonsei University College of Medicine,
Republic of Korea

REVIEWED BY

Murat Akand,
University Hospitals Leuven, Belgium
Abdulghafour Halawani,
King Abdulaziz University, Saudi Arabia

*CORRESPONDENCE

Saram Abbas

✉ s.abbas11@newcastle.ac.uk

Kabita Adhikari

✉ kabita.adhikari@newcastle.ac.uk

†These authors share senior authorship

RECEIVED 10 October 2024

ACCEPTED 09 December 2024

PUBLISHED 07 January 2025

CITATION

Abbas S, Shafik R, Soomro N, Heer R and Adhikari K (2025) AI predicting recurrence in non-muscle-invasive bladder cancer: systematic review with study strengths and weaknesses.

Front. Oncol. 14:1509362.

doi: 10.3389/fonc.2024.1509362

COPYRIGHT

© 2025 Abbas, Shafik, Soomro, Heer and Adhikari. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

AI predicting recurrence in non-muscle-invasive bladder cancer: systematic review with study strengths and weaknesses

Saram Abbas^{1*}, Rishad Shafik¹, Naeem Soomro²,
Rakesh Heer^{3,4†} and Kabita Adhikari^{1*†}

¹School of Engineering, Newcastle University, Newcastle Upon Tyne, United Kingdom, ²Department of Urology, Freeman Hospital, Newcastle upon Tyne, United Kingdom, ³Division of Surgery, Imperial College London, London, United Kingdom, ⁴Centre for Cancer, Newcastle University, Newcastle upon Tyne, United Kingdom

Background: Non-muscle-invasive Bladder Cancer (NMIBC) is notorious for its high recurrence rate of 70–80%, imposing a significant human burden and making it one of the costliest cancers to manage. Current prediction tools for NMIBC recurrence rely on scoring systems that often overestimate risk and lack accuracy. Machine learning (ML) and artificial intelligence (AI) are transforming oncological urology by leveraging molecular and clinical data to enhance predictive precision.

Methods: This comprehensive review critically examines ML-based frameworks for predicting NMIBC recurrence. A systematic literature search was conducted, focusing on the statistical robustness and algorithmic efficacy of studies. These were categorised by data modalities (e.g., radiomics, clinical, histopathological, genomic) and types of ML models, such as neural networks, deep learning, and random forests. Each study was analysed for strengths, weaknesses, performance metrics, and limitations, with emphasis on generalisability, interpretability, and cost-effectiveness.

Results: ML algorithms demonstrate significant potential, with neural networks achieving accuracies of 65–97.5%, particularly with multi-modal datasets, and support vector machines averaging around 75%. Models combining multiple data types consistently outperformed single-modality approaches. However, challenges include limited generalisability due to small datasets and the “black-box” nature of advanced models. Efforts to enhance explainability, such as SHapley Additive ExPlanations (SHAP), show promise but require refinement for clinical use.

Conclusion: This review illuminates the nuances, complexities and contexts that influence the real-world advancement and adoption of these AI-driven techniques in precision oncology. It equips researchers with a deeper understanding of the intricacies of the ML algorithms employed. Actionable insights are provided for refining algorithms, optimising multimodal data utilisation, and bridging the gap between predictive accuracy and clinical

utility. This rigorous analysis serves as a roadmap to advance real-world AI applications in oncological care, highlighting the collaborative efforts and robust datasets necessary to translate these advancements into tangible benefits for patient management.

KEYWORDS

artificial intelligence, non-muscle-invasive bladder cancer, NMIBC, machine learning, recurrence, prediction

1 Introduction

Bladder cancer continues to be a significant health concern. Particularly in the UK, where it stands as the 11th most common cancer, necessitating efficacious diagnostic and management strategies to curtail its impact (1). With 28 new reported cases and 15 deaths daily, the disease's impact is undeniable (1). A disease of heterogeneous nature, bladder cancer is primarily categorized into two main types: non-muscle-invasive bladder cancer (NMIBC) and muscle-invasive bladder cancer (MIBC). Among them, NMIBC is more common, yet carries a risk of escalating into MIBC if left untreated or poorly managed. About 45% of untreated high-grade NMIBC cases escalate into MIBC (2–4). An alarming feature of NMIBC is its high recurrence rate post-treatment, reported to be 70–80% (5), requiring frequent monitoring and interventions.

Due to its high prevalence and recurrence rate, bladder cancer happens to be one of the costliest cancers to manage – it cost the EU approximately 4.9 billion in 2012 to treat bladder cancer (6, 7). The high cost of managing NMIBC is due in part to the expensive diagnostic procedures and follow-up care required for patients with recurrent NMIBC. For example, cystoscopy, the gold standard for monitoring, is expensive (£240–£2000 per visit) and invasive, contributing significantly to overall costs and involving indirect costs from lost productivity and risks like urinary tract infections (8, 9). From 2015–2022, flexible cystoscopy cost the NHS over £810 million in total (inflation-adjusted) (10–15). Guidelines recommend cystoscopy every 3–6 months for the first 2 years, then annually, leading to significant cumulative costs. The average 3-year cost per NMIBC patient in the UK was estimated at £8735, with annual costs ranging from £1218 for grade 1 recurrence cases to £3957 for grade 3 (16). These escalating costs juxtapose the financial burden of NMIBC with the pressing need for precise diagnostic techniques and reliable predictive tools that can reduce the frequency of costly cystoscopic follow-ups.

Traditionally, prognostication and risk assessment in bladder cancer have relied on tools like the American Joint Committee on Cancer TNM staging system (17, 18). The TNM system categorizes bladder cancer based on tumor size and invasion (T), lymph node involvement (N), and distant metastasis (M). While validated and

widely used, these systems lack the comprehensive integration of factors needed for precise prognostication (19). They do not encompass the full scope of factors necessary for precise prognostication. They seem to fall short when compared to predictive models that incorporate numerous clinical variables (19). Additionally, their design does not readily permit the integration of novel information such as molecular markers or complex bioinformatics data, which are becoming increasingly relevant in the era of personalized medicine (20). A multimodal approach is recommended which combines histopathological markers with the results of imaging studies.

Despite numerous published studies on NMIBC recurrence prediction models, their adoption in clinical practice remains surprisingly low. Reasons for this can range from the lack of demonstrated improvement in clinical decision-making upon external validation to the logistical difficulty of integrating these prediction tools into electronic medical records to be readily available to the physician at the point of care. This systematic review addresses these gaps by analyzing AI-based prediction models for bladder cancer recurrence, emphasizing the need for better validation and integration into clinical workflows.

The aim of this systematic review is to evaluate the performance and utility of AI-based predictive models in NMIBC recurrence. The research question, framed using the PICOS criteria, is as follows: *How do AI-based predictive models perform compared to traditional methods in forecasting recurrence in NMIBC patients?* Specifically:

- Population: Patients diagnosed with NMIBC.
- Intervention: AI-based predictive models.
- Comparison: Traditional statistical or clinical models.
- Outcome: Predictive performance metrics (e.g., accuracy, discrimination, AUC) and clinical utility.
- Study Design: Observational, retrospective, and prospective studies.

By synthesizing the findings of existing studies, this review aims to identify gaps in current methodologies, provide insights into effective AI-based approaches, and offer guidance for future research to enhance personalized and efficient bladder cancer management.

1.1 Non-muscle invasive bladder cancer

Bladder cancer is a disease characterized by the uncontrolled growth of abnormal cells in the lining of the bladder, a hollow organ in the lower abdomen responsible for storing urine. Non-muscle-invasive bladder cancer (NMIBC) is a subtype of bladder cancer that hasn't penetrated the muscular wall of the bladder. This type of bladder cancer, often found early, comprises about 75% of all bladder cancer cases (21).

Bladder cancer staging follows the TNM system, detailing cancer progression within or beyond the bladder (see Figure 1 for TNM staging). Early stages (CIS, Ta, T1) are Non-Muscle-Invasive, with the tumor confined to the bladder's surface or connective tissue yet not penetrating the deeper muscle layers. These stages can, however, progress to muscle-invasive (MIBC) stages. NMIBC carries a lower risk of metastasis compared to MIBC but has a high recurrence rate (21). NMIBC has a tendency to often recur. These recurrences can be either at the same stage as the initial tumor or at a more advanced stage (22). Treatment primarily includes transurethral resection of bladder tumor (TURBT) to remove cancerous cells (23). Depending on the risk of recurrence and progression, intravesical therapy, where medication is directly instilled into the bladder, may be applied post-TURBT. This can include chemotherapy agents, such as mitomycin C, or immunotherapy with Bacillus Calmette-Guerin.

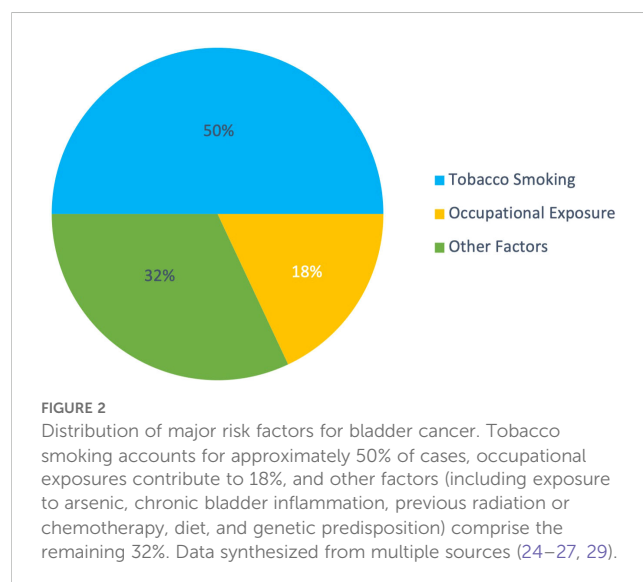
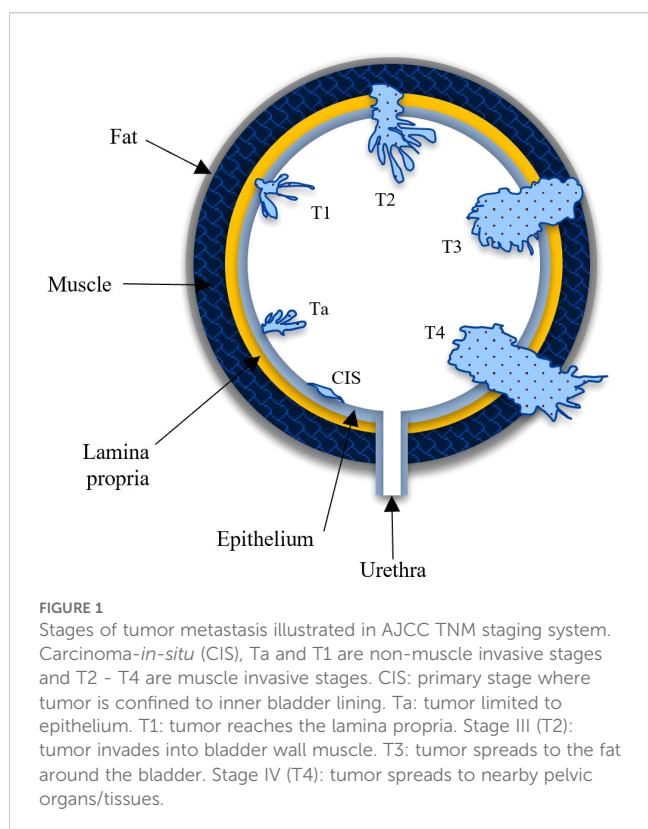
Bladder cancer is associated with numerous risk factors, the most significant being tobacco smoking, which accounts for about half of all cases (24). Figure 2 highlights the overwhelming effect of smoking over other factors. Other factors include occupational

exposures, chronic bladder inflammation, and genetic predisposition, which contribute to the complexity of predicting the disease's occurrence and progression (25–27). Additionally, gender disparities exist, with men being about three to four times more likely to develop bladder cancer than women (28). Age, diet high in processed meat, and tumor characteristics like grade and size also play crucial roles in predicting the likelihood of recurrence in NMIBC (24, 26, 29). Other predictive factors include genetic alterations or mutations, the presence of bladder cancer markers in the urine, and findings from imaging studies and cystoscopy (30). This variability in risk factors highlights the challenges in developing personalized care plans, underscoring the need for sophisticated predictive tools.

Creating personalized care plans for bladder cancer is challenging due to its highly varied nature. Each case differs in terms of tumor biology, stage, grade, and genetic mutations, which means a treatment effective for one patient may not work for another. Despite advances in bladder cancer genetics, the lack of reliable predictive biomarkers underscores the need for more advanced tools. In the following sections, we explore how Artificial Intelligence and Machine Learning (ML)-based approaches, are poised to fill this gap by offering more precise and individualized predictions, ultimately improving patient outcomes and treatment strategies.

1.2 Current methods and the need for innovation

The European Organisation for Research and Treatment of Cancer (EORTC) and The Club Urológico Español de Tratamiento Oncológico (CUETO), are two popular clinical urology research organisations, that developed the tools for recurrence prediction used currently. EORTC developed the EORTC Risk Tables (31), and CUETO introduced the CUETO Scoring system. These tables use a variety of factors, such as the number and size of tumors, prior recurrence rates, T-stage, and grade to calculate risk scores (32, 33).



Vedder et al. (34) conducted a study which revealed that both EORTC and CUETO's risk scores were not accurate in predicting recurrence (found *c*-indices of 0.55-0.61, where 0.5 is random guess). Other studies have also concluded that CUETO and EORTC presented poor discriminative value in predicting clinical events. These models overestimated the risk, especially in highest-risk patients (35–37). This has significant implications for both physicians and patients alike.

Poor and unreliable recurrence prediction has far-reaching consequences. Unreliable predictions of recurrence can trigger unnecessary invasive procedures like repeated cystoscopies or biopsies, overburdening healthcare providers and depleting valuable resources (38). This, in turn, escalates the financial costs borne by the NHS (39, 40), diverting funds that could otherwise be allocated to patient care and research endeavors. For patients, the repercussions of unreliable recurrence prediction are substantial, entailing risks and complications. False-positive predictions cause unnecessary anxiety and subject patients to additional tests and treatments with inherent risks. Conversely, false-negative predictions can delay or miss disease progression, hindering timely intervention and adversely affecting patient health (41, 42).

Traditional mathematical and statistical tools use a limited set of variables. These tools assume straightforward, proportional relationships between patient variables like tumor size, number of tumors, prior recurrence, and T-stage with outcomes such as recurrence and progression. By assigning linear scores based on cancer grade (1, 2, or 3), they overlook the nuanced severity differences, such as the greater jump from Grade 2 to 3 compared to Grade 1 to 2. In reality, cancer prognosis is shaped by complex, non-linear interactions involving molecular markers, genetic factors, and evolving medical data—factors that static, linear models cannot adequately capture. These traditional methods also struggle with missing data and incorporating recent advancements. Emerging approaches aim to enhance tumor classification, discover novel biomarkers, and improve predictions for bladder cancer metastasis, while other promising areas include body composition analysis and biomarkers beyond traditional assessments (43–45). Additionally, novel prognostic models based on gene signatures show potential (46). However, these innovations cannot be integrated into traditional tools highlighting the urgent need for more sophisticated models that offer reliable, dynamic predictions for NMIBC recurrence.

1.3 AI as a game-changer in NMIBC recurrence prediction

Artificial Intelligences (AI) has emerged as a powerful tool in the medical field, with machine learning (ML) at its forefront, particularly in tasks related to cancer prognosis and recurrence prediction. AI, a broad category of computational methodologies designed to emulate human cognitive functions, has been increasingly deployed in the medical field. ML algorithms can be trained to learn from existing data, adjusting their mathematical parameters to predict outcomes with high accuracy. In the context of bladder cancer, algorithms like support vector machines (SVM),

random forest (RF), artificial neural networks (ANN), and deep learning (DL) have been used to design models that enhance the prediction of cancer recurrence (47–52).

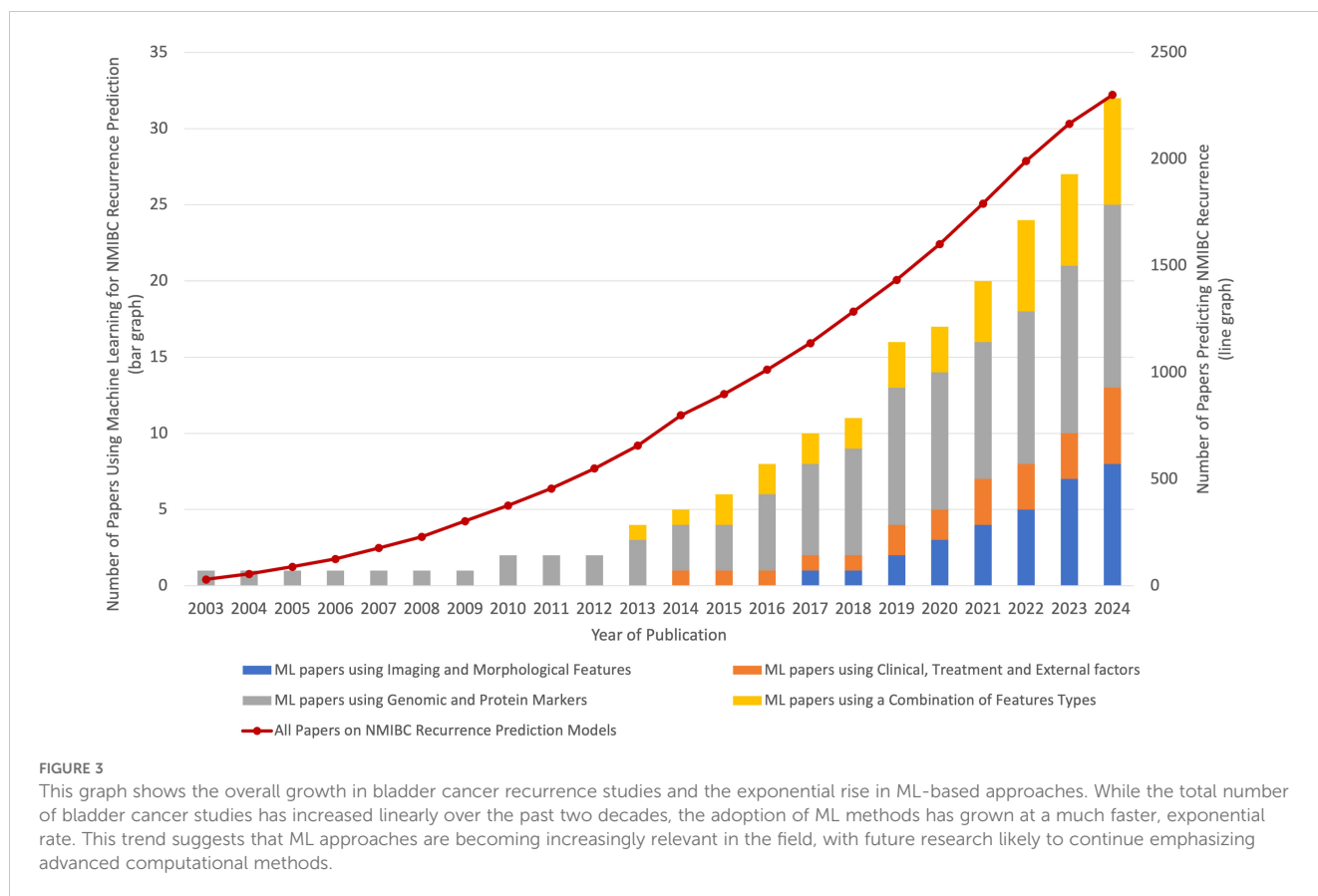
These models leverage a diverse range of markers — radiomic, clinical, pathological, and genomic — to build comprehensive and nuanced predictive models (53–57). They aim to improve risk stratification, anticipate recurrence, and optimize treatment planning, opening doors to more personalized patient management. Nonetheless, while advancements in AI and ML hold considerable promise, there is still much to learn, explore, and validate before these technologies can fully realize their potential in bladder cancer management.

Research in AI techniques to predict NMIBC recurrence is accelerating rapidly. As demonstrated in Figure 3, the number of studies focusing on bladder cancer recurrence has shown a steady, linear increase over the past two decades (depicted by a line graph). However, the adoption of ML approaches has grown at a much faster, exponential rate (depicted by the bar graph). This surge in ML research highlights the recognition of its potential to improve predictive accuracy and model complex, non-linear relationships that traditional statistical methods often struggle to capture. With increasing accessibility and advancements in computational power, the relevance of ML approaches in bladder cancer research is expected to continue expanding.

An emerging trend in the development of ML models for NMIBC prediction is the integration of multiple feature types, such as clinical, genomic, and imaging data. Single-feature models often fail to account for the multifaceted nature of cancer progression, whereas combining multiple data sources allows for more comprehensive models and improved predictive performance. This shift can be attributed to advancements in data integration techniques, the growing availability of multi-modal datasets and ease of access of powerful GPUs, which enable researchers to leverage richer and more diverse data for more accurate and clinically relevant predictions.

Figure 4 depicts a typical bladder cancer prediction workflow using an ML model, consisting of data pre-processing and ML algorithm application. The process begins with secure data storage, followed by data pre-processing, where irrelevant or incomplete data is removed and medical scans are segmented to focus on key areas like tumor lesions. Redundant features are identified and eliminated, often using clustering techniques. Then, ML algorithms are trained with this processed data for cancer recurrence prediction. The model's performance, evaluated by accuracy, precision, recall, and F1-score, may lead to retraining for improved standards. While accuracy measures the overall correctness of the model, precision indicates the proportion of positive cases correctly identified, avoiding false positives. Similarly, recall reflects how many actual positives were correctly identified, avoiding false negatives. The F1-score combines precision and recall into a single metric, balancing both false positives and false negatives. Finally, the model's effectiveness is compared with existing methods to assess any advancements in bladder cancer recurrence prediction and clinical detection benefits.

Accurately assessing the risk and predicting the occurrence and recurrence of NMIBC early is crucial for effective treatment and



management of NMIBC. ML-based diagnosis and predictive systems have been proven valuable in tumor detection, bladder segmentation, and NMIBC identification (48, 58). Moreover, AI approaches also have been applied to tumor staging, grading, survival rate prediction, response to chemotherapy, and recurrence rates, all of which are essential in personalized NMIBC management (59–61).

However, as of this date, no reliable AI algorithm is capable of accurately predicting NMIBC recurrence and improving management through a combination of the aforementioned markers. By reviewing these studies, we aim to pave the way for researchers to develop highly precise AI-based NMIBC recurrence prediction systems, facilitating optimal personalized management for early-stage NMIBC.

2 Methods

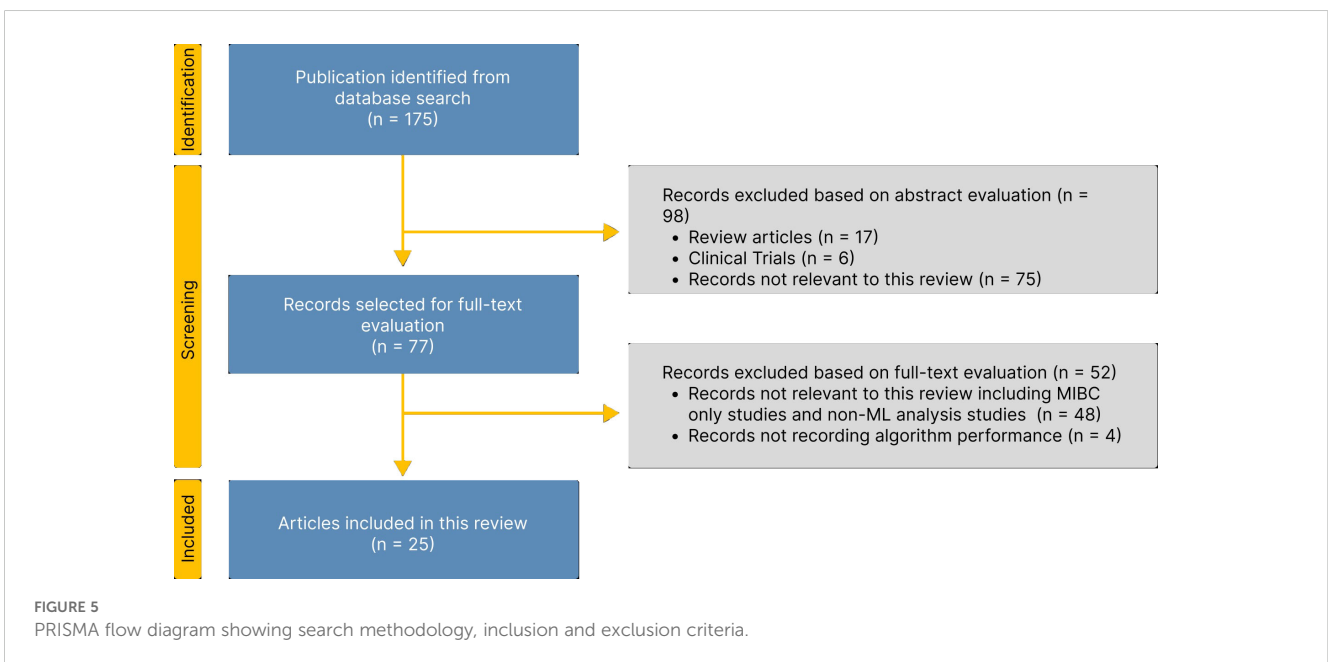
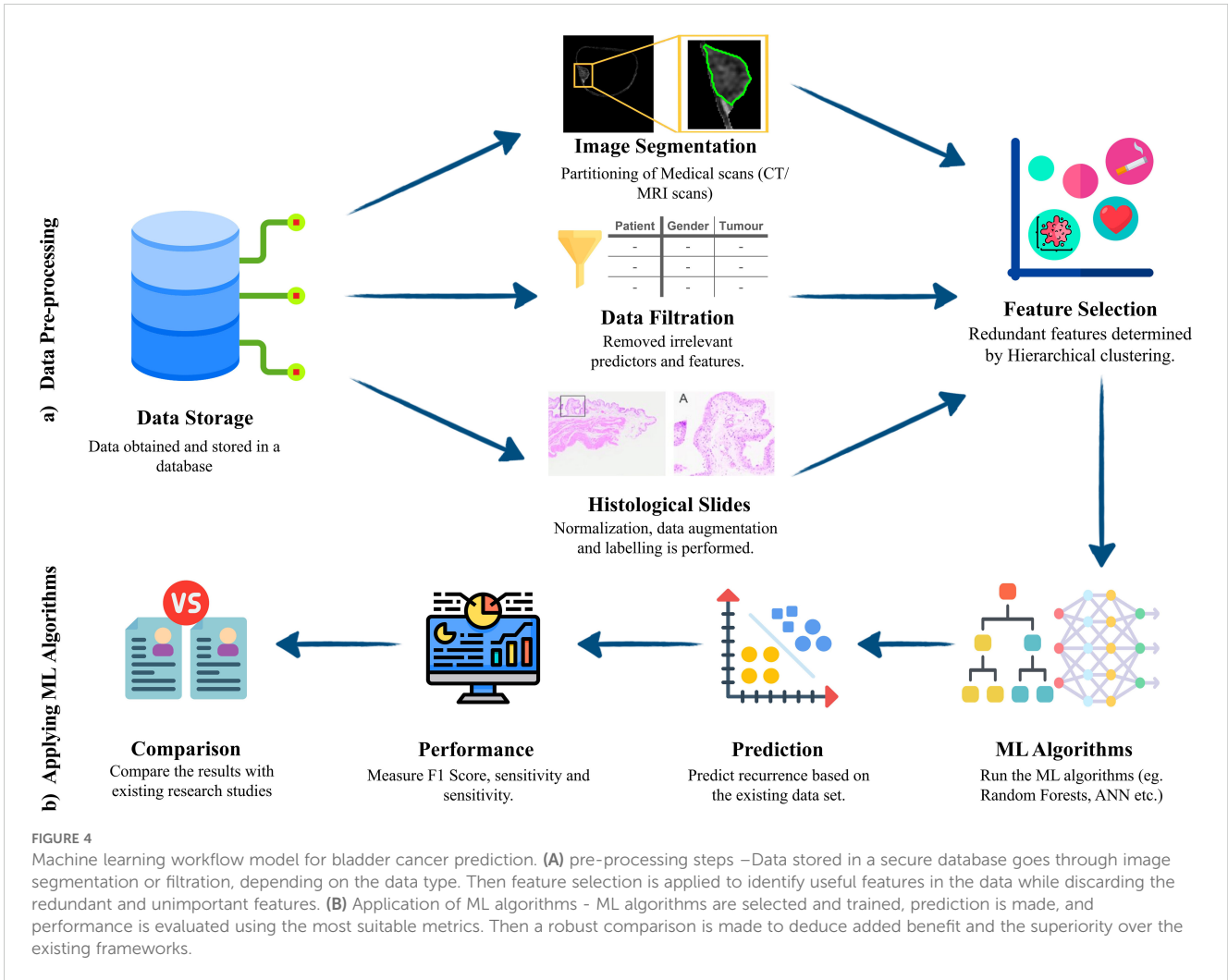
2.1 Search strategy

In this systematic review, we survey studies published up to October 2024, focusing on highly-ranked articles related to the implementation of ML and AI in bladder cancer prediction. We searched databases such as PubMed, IEEE Xplore, ScienceDirect (Elsevier), Springer, Nature, and MDPI. We used a combination of keywords that matched the scope of the survey such as “Bladder Cancer OR Non-Muscle Invasive Bladder Cancer OR NMIBC”

AND “Artificial Intelligence” OR “Machine Learning” OR “Neural Networks” AND “Prediction” OR “Predictor” AND “Recurrence”. Our search yielded a total of 175 unique studies. We reviewed the 175 articles on ML for NMIBC, focusing on those used for recurrence prediction. Upon abstract evaluation, 98 articles were discarded according to the exclusion criteria defined below. A further 46 studies were excluded based on full-text analysis. A total of 25 studies were finally selected for in-depth analysis within this review. Our study adhered to the guidelines set forth by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (62). The PRISMA flowchart is depicted in Figure 5.

2.2 Inclusion criteria

The inclusion criteria for this systematic review were defined to ensure the relevance and quality of the studies selected for analysis. The criteria are as follows: (1) original research articles, including peer-reviewed journal papers and conference proceedings; (2) studies specifically focused on the application of ML techniques for the prediction of NMIBC recurrence; (3) research exclusively addressing NMIBC recurrence or survival outcomes; (4) publications available in the English language. (5) Articles that report specific performance metrics (e.g., accuracy, sensitivity, specificity, AUC) to evaluate the effectiveness of the proposed ML models.



2.3 Exclusion criteria

The exclusion criteria for this systematic review were designed to eliminate studies outside the scope of NMIBC recurrence prediction. The following were excluded: (1) studies focused solely on diagnosis without addressing recurrence or survival prediction; (2) non-original research, including systematic reviews, meta-analyses, commentaries, editorials, and case reports; (3) articles lacking full-text access or sufficient methodological details; (4) studies using only traditional statistical methods without machine learning; (5) research on cancer types other than NMIBC (such as Gall Bladder Cancer, prostate cancer, lung cancer) or studies that did not differentiate between muscle-invasive and non-muscle-invasive bladder cancer.

3 Comparative analysis of AI models for NMIBC recurrence prediction

Accurate prediction of bladder cancer recurrence is critical for guiding treatment strategies and optimizing patient care. Recent breakthroughs in AI have ushered a new era of possibilities, offering unprecedented opportunities to refine and improve the prediction of recurrence as well as the occurrence of various fatal diseases such as cancers. By harnessing advanced ML algorithms and integrating diverse markers such as radiomic, clinical, histopathological, and genomic data, AI-based approaches have demonstrated remarkable potential in unraveling the intricate nature of bladder cancer recurrence. In this section, we explore the pivotal studies and advancements that showcase the transformative role of AI in bladder cancer recurrence prediction, shedding light on its promises, challenges, and future prospects.

To provide a clear understanding and a structured walkthrough of the studies, this section is categorized into four subsections based on similar methodologies as follows:

- Models using Imaging and Morphological Features
- Models using Genomic and Protein Markers
- Models using Clinical, Treatment and External Factors
- Models using a Combination of Feature Types listed above

To provide a consolidated view of research in this interdisciplinary field, we have provided the [Table 1](#) which offers a comprehensive overview of the various studies that have used ML techniques for NMIBC recurrence prediction. We provide the primary objectives of each study, the specific prediction tasks undertaken, the patient cohorts involved, the variables and data modalities employed, the chosen ML models, and the resultant performance metric.

3.1 Models using imaging and morphological features

Imaging-based ML models have become central to predicting NMIBC recurrence, leveraging data from pathology slides,

radiomics, and CT scans. These approaches provide detailed insights into tumor morphology and cellular patterns, often revealing features that traditional methods may overlook. Studies in this category use deep learning and other ML techniques to enhance prognostic accuracy by analyzing both tissue samples and imaging modalities.

The analysis of pathology slides has been a cornerstone of many ML models for NMIBC recurrence prediction, offering insights into tissue morphology and cellular characteristics. Chen et al. (63) used machine learning on Hematoxylin and Eosin-stained images from 514 patients, a substantial sample size for bladder cancer research. Unlike other approaches that use nuclear extraction techniques, they applied LASSO with 10-fold cross-validation to identify 22 bladder cancer-related and 18 survival-related image features. Their diagnostic model achieved strong performance, with Area Under the Receiver Operating Curve (AUROC) values of 94.1%, effectively distinguishing bladder cancer from normal tissues and glandular cystitis. The ML-based risk score served as an independent predictor for survival, enhancing prediction accuracy for 1-, 3-, and 5-year overall survival by more than 10%. However, the use of non-uniform median cut-off values for high-risk scores limits consistency, indicating a need for further standardization and validation in future prospective trials before broad clinical adoption.

Building on this approach of using pathology slides, Tokuyama et al. (49) explored the predictive value of nuclear atypia in NMIBC, focusing on the morphological characteristics of cancer cells derived from transurethral resection specimen. They utilized ML to predict the recurrence of NMIBC based on nuclear atypia (abnormalities in the nuclei of cancer cells) extracted from transurethral resection specimens. Using SVM and RF algorithms on a dataset of 125 patients, the authors derived quantitative morphological features from regions of interest on Hematoxylin and Eosin-stained slides. This involved the application of a nuclear extraction process using software programs “Ilastik” and “YOLO v3” for the segmentation of individual nuclei. The SVM-based model achieved a 90% probability of predicting NMIBC recurrence within 2 years post-TURBT, while the RF-based model achieved 86.7%. Despite these promising results, the study’s limitations included a relatively small sample size, the potential for bias in the nuclear extraction process, and the limited generalizability due to the exclusion criteria.

Drachneris et al. (64) provided a new perspective by analyzing immune cell density gradients. They used CD8+ cell-density gradients to predict Recurrence-free survival in NMIBC patients post-BCG therapy. CD8+ are a type of white blood cell that is crucial for fighting infections or cancer. Traditional methods merely count the total number of immune cells, while this study analyses the spatial distribution (or the cell density gradients), which provides deeper insight into the cancer’s immune response. By combining the HALO AI Densenet v2 classifier (a deep learning-based image analysis platform designed for digital pathology) and multivariable Cox regression models, the study creates a robust model with a C-index of 0.74. However, the paper recognizes that broader, prospective validations and explorations of other immune factors are required to further validate the methods. This study shows that a fresh perspective on established methods can enhance bladder cancer prognosis accuracy.

TABLE 1 Summary of studies utilizing machine learning techniques for NMIBC recurrence prediction and assessments.

Reference	Objective	Prediction Task	Patients	Variables	Data Modality	MLModel	Performance
Tokuyama et al. (49)	Predict NMIBC recurrence	Recurrence	125	Nuclear atypia from TUR specimens	H&E stained slides	SVM, Random Forests	SVM: 90.0%, RF: 86.7%
Chen et al. (63)	Diagnostic and prognostic models for BCa based on H&E images	Diagnosis and survival	514	H&E stained images	Digital pathological images	LASSO-Cox hazard model	Diagnostic model: AUROC: 89.2% - 96.3%; Prognostic model: HR = 2.09 (TCGA cohort), HR = 5.32 (General cohort)
Xu et al. (47)	personalized tool for BCa recurrence risk	Recurrence risk	71	Radiomics and clinical factors	MRI and clinical data	SVM with feature elimination	Accuracy: 81.0% AUROC: 83.8%
Levy et al. (65)	Predict bladder cancer recurrence using imaging features from urine cytology	Recurrence	135	Imaging features from urine cytology exams	Urine cytology slide images	AutoParis-X (NN)	C-index: Up to 0.77
Huang et al. (67)	Predict 5-year recurrence risk in NMIBC using MRI and deep learning	Recurrence	191	Clinical, radiomics, and deep learning features	Multiparametric MRI	Clinical and Radiomics	AUC: 0.909, C-index: 0.804
Wang et al. (66)	Predict recurrence of NMIBC using deep learning on pathology images	Recurrence	210	Pathological and clinical features	Pathology slide images	Deep learning (patch-level, WSI-level)	AUC: 0.860
Lee et al. (69)	Significance of IPP in NMIBC prognosis	Prognosis	122	IPP, age, BPH and other clinicopathological variables	Clinical data	SVM	Accuracy: 80.0%
Zhao et al. (70)	Predict prognosis and recurrence of bladder cancer using ECM-related genes	Prognosis and Recurrence	960	Six ECM-related genes, FSTL1, stage, age, gender	Genetic and clinical data	GLM, KNN, SVM, Random Forests	AUC: 76%
Cai et al. (71)	Evaluate prognostic role of LOH on chromosome 18 in low-risk NMIBC	Recurrence	108	LOH on chromosome 18, number of lesions, clinico-pathological factors	Clinical data	Univariate and multivariate analyses, NN analysis	Multivariate analysis found LOH as an independent predictor of recurrence-free probability; NN performance metrics not mentioned.
Wang et al. (68)	Predict bladder cancer recurrence using multiphase CT images	Recurrence	874	Clinical and imaging features	Multiphase enhanced CT images	Convolutional Neural Networks	AUC: 0.889, C-index: 0.869
Urbanowics et al. (76)	Identify genetic and environmental factors in bladder cancer susceptibility	Risk and Survival time	914	DNA repair gene SNPs, smoking history	Genetic and clinical data	AF-UCS (Learning Classifier System)	Accuracy: 66%
Maturana et al. (82)	SNP impact on NMIBC prognosticators	Prognosis	995	171,304 SNP, 6 clinical-pathological indicators	Clinical and genetic data	Bayesian learning with LASSO	Time to First Recurrence AUC: Clinico-pathological = 0.62 SNP-only = 0.55 Combined = 0.61. Time to progression AUC: Clinico-pathological = 0.76

(Continued)

TABLE 1 Continued

Reference	Objective	Prediction Task	Patients	Variables	Data Modality	MLModel	Performance
							SNP-only = 0.58 Combined = 0.76
Krochmal et al. (85)	Predict bladder cancer recurrence using Urinary peptide panel	Recurrence	98	Urinary peptide panel	Urine samples	Random Forests	AUC: 0.828
Frantzi et al. (84)	Develop and validate urine-based biomarker panels for primary and recurrent BCa detection	Biomarker panels for detection	481	Peptide biomarkers	Urine samples		Primary panel: Accuracy 76%, AUC 0.77
Zhan et al. (86)	Develop a urinary exosome derived lncRNA panel for BCa diagnosis and recurrence	Diagnosis and Recurrence	368	MALAT1, PCAT-1, SPRY4-IT1 lncRNA expressions	lncRNA data	Multivariate Logistic Regression	AUC: 0.813 (validation)
Gogalic et al. (87)	Validate a protein panel for noninvasive detection of recurrent NMIBC	Diagnosis and recurrence	45	Biomarkers (ECadh, IL8, MMP9, EN2, VEGF, past recurrences, BCG therapies, stage at diagnosis)	Urine samples	Logistic regression with LASSO	Highest AUC: 0.96
Zhang and Ma (89)	Identify biomarkers for bladder cancer recurrence using ML	Recurrence	345	TBIL, CA50, and 34 clinical parameters	Clinical and biomarker data	Decision tree, Random Forest, AdaBoost, GBM, XGBoost	Best model AUC: 0.623
Chang et al. (83)	Develop non-invasive urine-based proteomic biomarkers for recurrence	Recurrence	279	11 Urinary proteins	Urine proteomic data	mSVM-RFE	AUC: 0.784
Mucaki et al. (61)	Predicting chemotherapy responses	Recurrence and remission	54	Gene expression, clinical metadata	RNA-seq, microarray	SVM	Recurrence Acc: 71.0% (cisplatin), 60.2% (carboplatin), 54.5% (oxaliplatin) Remission Acc: 59%, 61%, and 72%
Drachneris et al. (64)	Predict recurrence-free survival in NMIBC post BCG therapy	Recurrence-free survival	157	CD8+ cell density indicators (Immunodrop, Center of Mass), tumor stage, tumor grade	Histopathological slides	Multivariable Cox regression	C-index: 0.7837
Ajili et al. (88)	Predict BCa recurrence post-BCG immunotherapy	Recurrence classification	308	Clinical variables: Age, gender, tumor stage, grade, size, multiplicity, smoking, CD34 expression	Histopathological data	Neural Networks	Sensitivity: 96.66%, Specificity: 100%, PPV: 100%, NPV: 90.9%
Catto et al. (52)	To compare the predictive accuracies of NFM, NN, And traditional statistical methods for the behavior of BCa	Recurrence (Occurrence and Timing)	109	Experimental molecular biomarkers (p53, mismatch repair proteins) and conventional clinicopathological data	Clinical and molecular data	NFM, Neural Networks, Logistic Regression and Linear Regression	NFM and NN Accuracy: 88-95% LR and LinR accuracy = 71-77%

(Continued)

TABLE 1 Continued

Reference	Objective	Prediction Task	Patients	Variables	Data Modality	MLModel	Performance
Lucas et al. (91)	Predict NMIBC recurrence	1- and 5-year RFS	359	Digital histopathology slide data and clinical data	Digital histopathology slides and clinical records	Convolutional Neural Networks	AUC: 0.62 (1-year) and 0.76 (5-year)
Jobczyk et al. (92)	Recalibrate prediction tools for the recurrence and progression of NMIBC	RFS and PFS prediction	3,892	Gender, age, T stage, histopathological grading, tumor burden and diameter, EORTC and CUETO scores, intravesical treatment	Clinical data	DeepSurv (Deep Neural Networks)	RFS: C-index 0.65, PFS: C-index 0.88
Schwarz et al. (90)	Improve explainability in ML for predicting UBC recurrence	Recurrence	1,944	Clinical and therapeutic features	Tabular clinical data	SVM, Gradient Boosting, ANN	AUC: 70.82%, F1-Score: 83.89

ACCI, Age-adjusted Charlson Comorbidity Index; ANN, Artificial Neural Networks; ASA, American Society of Anesthesiologists; BCa, Bladder Cancer; BPN, Back Propagation Neural network; CCI, Charlson Comorbidity Index; CHF, Congestive Heart Failure; DL, Deep Learning; DSS, Disease-Specific Survival; ELM, Extreme Learning Machine; EMR, Electrical Medical Records; GBT, Gradient Boosting Trees; HR, Hazard Ratio; IPP, Intravesical Prostate Protrusion; KNN, K-Nearest Neighbors; LR, Logistic Regression; LinR, Linear Regression; MLP, Multilayer Perceptron; NB, Naive Bayes; NBI, Narrow Band Imaging; NFM, Neuro-fuzzy modeling; NMIBC, Non-Muscle Invasive Bladder Cancer; NPV, Negative Predictive Value; OR, Odds Ratio; OS, Overall Survival; PFS, Progression-Free Survival; PPV, Positive Predictive Value; RBFN, Radial Basis Function Network; RELM, Regularized Extreme Learning Machine; RF, Random Forest; RFS, Recurrence-Free Survival; SNP, Single Nucleotide Polymorphism; SVM, Support Vector Machine; TCGA, The Cancer Genomic Atlas; TUR, Transurethral Resection; UE, Urinary Exosome; WL, White Light.

The use of cellular features extends to urine cytology as well. The study by Levy et al. (65) used a Deep learning ML tool called AutoParis-X to predict bladder cancer recurrence risk from 1,259 urine cytology images. The prediction model achieved good accuracy (C-index 0.77), outperforming models using standard cytological assessment alone. It was found that the model worked best when it looked at samples from the first 6 months after the initial cancer diagnosis. This model was trained on retrospective data from 159 patients from only one hospital. The small number of patients, single-centered and retrospective analysis results in concerns about the model’s generalizability.

Pathomics, which integrates patch-level and whole-slide image analyses, represents another advancement in pathology-based models. Wang et al. (66) introduced a novel pathomics model using deep learning for predicting NMIBC recurrence with high accuracy. This study aimed to address the challenge of early recurrence prediction in NMIBC by leveraging deep learning to analyze pathology images. The model, developed using a two-phase approach—patch-level prediction followed by whole slide image (WSI)-level prediction—achieved a strong performance with an AUC of 0.860 in the test cohort. Transfer learning was employed to generalize the model across different datasets, while model interpretability was improved through visualization techniques, helping clinicians understand the predictions. However, the use of pathology images, while innovative, requires digital pathology infrastructure that may not be available in all clinical settings. The model’s excellent performance suggests it could be a valuable tool in clinical practice, but external validation in larger and more diverse patient populations is needed.

Shifting from pathology images to radiomics, Xu et al. (47) shifted the focus toward integrating radiomics features from MRI scans, offering a novel approach to personalized risk assessment. Xu et al. conducted a study which aimed to devise a personalized tool for estimating the two-year recurrence risk of bladder cancer. A model incorporating both radiomics features extracted from MRI scans and clinical factors was built using data from 71 patients. The model utilized ML methods like SVM-based recursive feature elimination. This nomogram, relying on muscle-invasive status and an radiomics-derived score, showed promising accuracy (80.95% in validation) and a high AUROC value of 0.838. Despite these strong results, potential bias due to retrospective design and single-center sampling remain as the limitations. Future investigations could explore the potential of additional factors, currently omitted due to incomplete data, in enhancing the predictive power of the model. Moreover, the role of different radiomics features in predicting lymph node status, crucial for prognosis, was recommended for future research.

Building on this, Huang et al. (67) demonstrated the potential of combining multiparametric MRI and Deep learning for predicting bladder cancer recurrence. In their study, the authors developed a clinical-radiomics deep learning model that integrated radiomics features from multiparametric MRI with deep learning and clinical data to predict the 5-year recurrence risk in NMIBC patients. With a dataset of 191 patients, the model outperformed traditional clinical models, achieving an AUC of 0.909. The use of SHapley Additive ExPlanations further enhanced the interpretability of the

model, showing that radiomics features contributed significantly to prediction accuracy. However, the approach relies heavily on expensive and advanced MRI technology, which may not be accessible in many clinical settings. Furthermore, while SHAP values increase interpretability, the inherent complexity of radiomics and the deep Learning models still creates a “black box” issue, limiting clinician trust in the model’s outputs.

CT imaging has also been explored for recurrence prediction. Wang et al. (68) developed a deep learning model using multiphase enhanced CT images to predict bladder cancer recurrence, showing substantial clinical promise. This multi-center study involved 874 patients from four centers and used CNNs to develop a signature capable of predicting recurrence risk. The model demonstrated excellent performance, with an AUC of 0.889 and a concordance index of 0.869. It outperformed traditional clinical models and staging systems, indicating its potential utility in guiding personalized treatment strategies for NMIBC patients. Despite the robust performance of the deep learning model, it relies on high-resolution imaging data, which may limit accessibility in less technologically advanced medical centers. Further prospective validation and exploration of how the deep learning model integrates with clinical decision-making would be beneficial.

Finally, specialized clinical features, such as intravesical prostatic protrusion (IPP), have been studied as independent predictors of recurrence. Lee et al. (69) investigated the role of IPP in predicting NMIBC recurrence using imaging data obtained from preoperative CT urography. By analyzing the severity of IPP in 122 male NMIBC patients, the study demonstrated that structural bladder and prostate features significantly influence recurrence risk. Severe IPP ($\geq 5\text{mm}$) was observed in 27% of patients and was associated with a 2.6-fold increased risk of recurrence. Using SVM, incorporating IPP improved NMIBC recurrence prediction by 6%. Kaplan-Meier analysis showed that severe IPP negatively impacted recurrence-free especially in high-risk patients. However, limitations such as the small sample size, retrospective design, and lack of post-void residual data suggest further research is needed to explore the relationship between IPP, ageing, and bladder conditions. These findings position IPP as an independent risk factor for recurrence and its potential role in prognosis.

3.2 Models using genomic and protein markers

A multitude of studies have attempted to predict NMIBC recurrence, each employing unique yet complementary methodologies using biological insights combined with ML. The journey begins with Zhao et al.’s (70) paper, which focuses on ECM-related genes’ prognostic potential in Bladder Cancer. By using expression data from multiple datasets, six ECM-related genes were identified: CTHRC1, MMP11, COL10A1, FSTL1, SULF1, COL5A3. An AUC of 0.76 was achieved in recurrence prediction by training many ML models, including Generalised Linear Models, K-Nearest Neighbours, SVMs and Random Forests, on 675 non-recurrent and 285 recurrent bladder cancer patients. The study demonstrates the

potential of ML and ECM gene signatures for recurrence prediction in bladder cancer, though larger sample sizes would help validate the approach. Limitations include few normal samples for comparison and lack of extensive validation of individual gene trends.

Another genomic study, this time by Cai et al. (71), analyzed loss of heterozygosity (LOH) on chromosome-18 in 65 patients with NMIBC and 43 controls. Loss of heterozygosity occurs when one of the two copies (alleles) of a gene, inherited from each parent, is lost or inactivated, which can lead to cancer if the affected gene’s role is to suppress tumors. At multivariate analysis, LOH on Chr 18 ($P=0.002$) and the number of lesions ($P=0.03$) were identified as independent predictors of recurrence-free probability. ANNs were used to confirm the multivariate analysis but the performance metrics were not mentioned. This paper, published in 2010, did not gain widespread adoption due to factors such as small study sample size; robustness of other recurrence predictors [FGR3 (72), Ki-67 (73) and NMP22 (74, 75)].

Urbanowicz (76) also aimed to uncover patterns of genetic associations with bladder cancer by applying an ML classifier system called AF-UCS (Attribute Feedback-supervised Classifier System). The algorithm aimed to validate findings that specific SNPs in DNA repair genes, such as XPD (Xeroderma pigmentosum group D) codon 751 and 312, along with SNPs in other DNA repair genes, are predictive of bladder cancer risk when considered alongside smoking pack-years. Their model had an accuracy of 0.66, indicating a low to moderate level of predictive performance. Since this publication in 2013, there has been limited follow-up potentially due to other ML methods having superior accuracy, and the other models not requiring specialized expertise. Additionally, large-scale Genome-Wide Association Studies are now more common than focused candidate gene approaches (77–81).

Building on the growing interest in genomic predictors, Maturana et al. (82) published a study three years later that focused on predicting outcomes of NMIBC patients using genomic SNP profiles in conjunction with clinico-pathological prognosticators. The study utilized Bayesian learning methods, including sequential threshold models and LASSO, and evaluated 822 NMIBC patients followed up for over a decade. The genomic models yielded AUROC values ranging from 0.55 to 0.62, while clinic-pathological models performed slightly better, with AUROC values between 0.57 and 0.76. These results indicate that SNP profiles alone are poor predictors of NMIBC recurrence and progression, and their inclusion in clinico-pathological models adds limited value. The limitations of the study include a relatively small sample size in certain subgroup analysis and the limited predictive ability of common SNPs in NMIBC outcomes.

Chang et al. (83) utilized urinary proteomic profiling to develop a non-invasive method for diagnosing and monitoring bladder cancer recurrence. By analyzing urine samples from 279 patients, the authors used a multi-support vector machine-recursive feature elimination (mSVM-RFE) algorithm to identify 13 protein markers for diagnosis and 11 markers for recurrence monitoring. The diagnostic model achieved high sensitivity (90.9%) and specificity (73.3%), while the recurrence monitoring model reached 75% sensitivity and 81.8% specificity. This study’s strength lies in its

non-invasive nature, providing a practical alternative to invasive procedures like cystoscopy. However, the reliance on high-resolution mass spectrometry and the single-center design may limit the generalizability of the findings. The model's utility is promising, but larger, multi-center validation studies are required to ensure its efficacy in diverse clinical settings.

Frantzi et al.'s study (84) formulated and validated urine-based biomarker panels for primary and recurrent bladder cancer detection using capillary electrophoresis-mass spectrometry (CE-MS). Their methodology Combined statistical analysis and machine learning: initial statistical tests identified significant peptide biomarkers, which were further refined using SVMs via MosaCluster software to create and optimize high-dimensional biomarker panels. Case-control comparisons across multicenter cohorts identified the biomarkers, with the primary panel achieving an AUROC of 0.87 (91% sensitivity, 68% specificity) and the recurrent panel attaining an AUROC of 0.75 (88% sensitivity, 51% specificity) during independent validation. By incorporating all available biomarkers, optimized panels improved performance further, achieving AUROCs of 0.88 for primary and 0.76 for recurrent cancer. These findings highlight the potential for non-invasive urine-based tests to complement or reduce the need for invasive cystoscopy in bladder cancer diagnosis and monitoring. Despite these promising results, the study was limited by unadjusted confounding variables, such as tumor size and hematuria, and by its cross-sectional design. Further research is needed to confirm their clinical utility and performance in real-world settings.

This study by Krochmal et al. (85) analyzed CE-MS peptidomics data (detected types and amounts of peptides in a biological sample) from 98 bladder cancer patients to develop an ML model to predict recurrence. With a training set of 50 patients, Cox regression identified 36 peptides predictive of relapse which were then input into a Random Forest model. The training set produced an accuracy of 100%, while test set accuracy is not mentioned. The limited size of the training dataset ($n=48$), and the unreported test accuracy, calls into question the study's generalizability. A comparison with established clinical tools would have highlighted any advancements over existing methods.

Mucaki et al. (61) applied biochemically-inspired machine learning (ML) models, specifically supervised SVMs, to predict responses to chemotherapy agents, including cisplatin, a critical drug in bladder cancer treatment. The models were described as "biochemically-inspired" because they incorporated genes with established biological relevance to the mechanisms of action and resistance of cisplatin, ensuring the selection of features grounded in prior biochemical knowledge rather than relying purely on data-driven approaches. These gene signatures were designed to reflect pathways involved in apoptosis, DNA repair, and drug transport, enhancing interpretability and biological validity. Using a dataset of 90 cancer patients, the cisplatin-specific gene signature achieved 71.2% accuracy in predicting bladder cancer recurrence, with exceptional performance in non-smokers (100%) and 79% accuracy in smokers. The study also utilized ensemble averaging across multiple thresholds to improve model robustness. However, a key limitation was the reliance on breast cancer cell line data to

train the SVM models, which limits their applicability to bladder cancer-specific contexts, as no retraining was performed on bladder cancer datasets. While this work highlights the promise of AI-driven approaches for identifying recurrence risk factors, it also underscores the need for developing bladder cancer-specific models to improve clinical relevance.

Two other studies [Zhan et al. (86) and Gogalic et al. (87)] explored the application of protein panels in recurrence prediction. Zhan et al. utilized MALAT1, PCAT-1, and SPRY4-IT1 biomarkers and achieved an accuracy of 81.3% and a sensitivity and specificity of 0.625 and 0.850 respectively. Additionally, tumor stage showed a statistically significant correlation as a predictor, with PCAT-1 identified as an independent predictor. Gogalic et al. combined common clinicopathological markers with ECadh, IL8, MMP9, EN2, and VEGF biomarkers. Their model, incorporating these markers, yielded an AUROC of 0.84.

3.3 Models using clinical, treatment and external factors

Ajili et al. (88) applied ANNs to predict bladder cancer recurrence after Bacillus Calmette-Guerin immunotherapy. Using a multilayer perceptron model in MATLAB, the researchers incorporated patient characteristics, tumor attributes, and treatment details, achieving high performance, with a mean square error of 0.02634. The model accurately classified 39 out of 40 cases, yielding sensitivity of 96.66%, specificity of 100%, and positive/negative predictive values of 100% and 90.9%, respectively. However, the small sample size and difficulties in determining the optimal network topology (e.g., hidden layer nodes) limited the study. These results showcase the potential of ANNs in bladder cancer prognosis but illuminate the need for larger datasets and further validation before clinical application.

Zhang and Ma (89) investigated the predictive value of two clinical biomarkers, CA50 and total bilirubin (TBIL), for bladder cancer recurrence using ML models. The study evaluated the individual and combined predictive performance of these biomarkers in a cohort of 345 bladder cancer patients. The results demonstrated that CA50 had an AUC of 0.602 ($p = 0.038$) and TBIL had an AUC of 0.585 ($p = 0.014$), indicating that both biomarkers were moderate predictors of recurrence. When combined, the AUC increased to 0.623 ($p = 0.013$), showing an improvement in predictive power, though still relatively modest. While the combination of these two biomarkers offered better performance than either one alone, the overall predictive accuracy remained limited, suggesting that additional features or biomarkers may be necessary to enhance the model's ability to predict bladder cancer recurrence. The study provides valuable insights into the role of clinical biomarkers in cancer prognosis but highlights the need for further optimization to improve predictive robustness.

Schwarz et al. (90) examined the role of explainability in ML models for predicting bladder cancer recurrence, enhancing their clinical utility. In this study, the authors used three ML models—SVM, gradient boosting, and ANNs—and compared them to logistic regression for predicting 2-year recurrence in urothelial

carcinoma patients. Gradient boosting performed best, with an F1-score of 83.89% and AUC of 70.82%. To address the black-box nature of these models, the authors employed permutation feature importance and feature importance ranking measure to explain the most influential features driving predictions, such as therapeutic measures. This approach enhances the transparency of ML models, potentially increasing their adoption in clinical settings. However, while they attempt to address the “black-box” issue using feature importance measures like permutation feature importance and feature importance ranking measure, the interpretability of these models remains limited. These feature importance methods often fail to provide clinically actionable insights, as they simply show which features contribute most to the prediction without explaining why.

3.4 Models using a combination of feature types

Integrating diverse feature types—such as imaging, morphological, clinical, and genomic markers—has proven to significantly enhance the predictive performance of AI models for NMIBC recurrence. By leveraging multiple data sources, these models offer a more comprehensive understanding of tumor behavior and patient risk profiles. This section reviews key studies that have successfully combined these feature types, highlighting their methodologies, outcomes, and potential for clinical application.

One of the earliest studies in this area, conducted by Catto et al. (52), aimed to compare the predictive accuracies of neuro-fuzzy modeling (NFM), neural networks (NN), and traditional statistical methods for predicting bladder cancer recurrence. The study used data from 109 NMIBC patients who were treated with TURBT. The data was used to train and test the predictive models. The study found that both NFM and NN predicted the patients' relapse with an accuracy ranging from 88% to 95%, which was superior to statistical methods (71-77%). The difference was statistically significant, indicated by the p-value of less than 0.0006. The low p-value suggests that these differences are not just due to random chance. NFM appeared better than NN at predicting the timing of relapse ($p < 0.073$). Importantly, NFM offered a transparency advantage over NN, allowing for easier clinical validation and manipulation of input variables for exploratory predictions. This early work highlighted the potential of combining different data types to boost predictive power while maintaining interpretability.

Lucas et al. (91) further advanced this approach by incorporating both histopathological image features and clinical data to predict recurrence-free survival in NMIBC patients. Their model, built using Convolutional Neural Networks (CNNs) and Bidirectional GRUs, achieved a 1-year AUC of 0.62 ($n=359$) and a 5-year AUC of 0.76 ($n=281$). Their prediction model consisted of a multi-step process that combined features extracted from the histopathological images with the clinical data to create an overall outcome. Their model's strength lies in the autonomous analysis of histopathological data, reducing human bias and possibly capturing more nuanced features. However, the study's relatively small,

single-center dataset and extended inclusion period raised concerns about generalizability and consistency in clinical practices over time. Nevertheless, the integration of image and clinical data illustrated the power of multi-modal approaches in predicting bladder cancer recurrence.

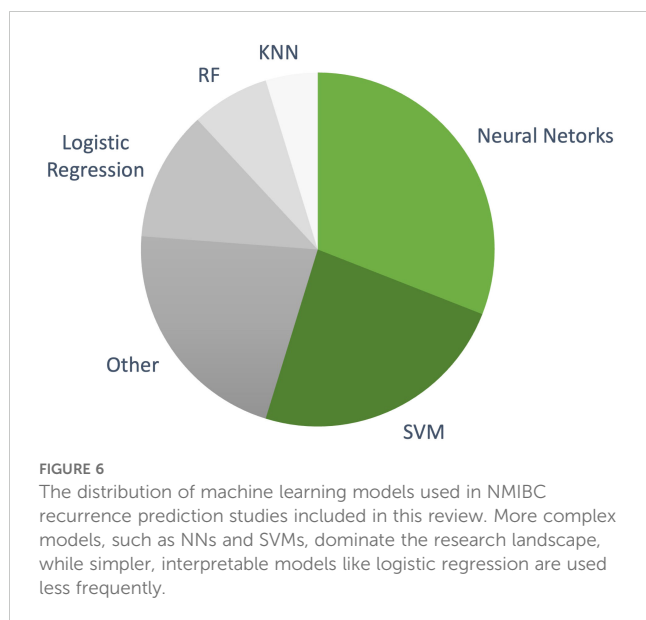
In this study (92), a deep learning-based approach (DeepSurv) was used by Jobczyk et al. to recalibrate the prediction tools for the recurrence and progression of NMIBC in a cohort of 3,892 patients. The existing risk groups, EORTC and CUETO, showed moderate performance in predicting survival outcomes. The deep learning models displayed improved accuracy. In the training group of 3,570 patients, the c-indices were 0.650 for recurrence-free survival and 0.878 for progression-free survival. In the validation group of 322 patients, the c-indices stood at 0.651 for recurrence-free survival and 0.881 for progression-free survival. The models surpassed the performance of standard risk stratification tools and demonstrated no signs of overfitting. These findings highlight the potential of deep learning models in enhancing the prediction of recurrence and progression in NMIBC, offering a valuable tool for personalized patient care.

4 Discussion

Our comprehensive review of 25 ML-based studies highlights a growing trend in the adoption of ML-based approaches for NMIBC prediction and management. This demonstrates the potential of these approaches to drive a transformative shift in bladder cancer care. Our work offers a distinct contribution as we move beyond surface-level summaries to conduct a rigorous, in-depth evaluation of each study's methodologies. We prioritize a detailed interpretation of the reported model performances, assess the robustness of sample sizes, explore the diversity of data modalities, and critically analyze the specific prediction tasks—key aspects that are often overlooked in other reviews. By focusing on these crucial factors, we uniquely uncover strengths and weaknesses that were previously missed, providing a deeper, more nuanced understanding of each study's true contribution to the field.

Our analysis reveals a significant evolution in the field. Early studies were confined to single data types, but as the field progressed, more advanced ML models integrating multiple modalities—such as radiomics from MRI and CT scans, genetic markers, and histopathology—showed considerable promise. This integration has led to superior predictive performance, emphasizing the importance of a holistic approach to NMIBC recurrence prediction. A significant number of studies on NMIBC recurrence prediction employ these complex ML models, as evidenced by Figure 6.

Among these, NNs have emerged as the most commonly used approach, featured in 10 studies (46, 52, 64–66, 68, 71, 88, 91, 92). These models, which incorporate clinical, pathological, and genomic markers, achieved accuracies ranging from 0.65 to 0.975. NNs are particularly suited to handling complex, multi-modal datasets and often outperform simpler models in predicting recurrence, making them an attractive option for integrating



various data types. SVMs were the second most used algorithm, appearing in 9 studies (47, 49, 61, 69, 70, 83, 85, 90, 92). These studies reported overall accuracies around 0.75. SVMs are known for handling high-dimensional data well, maintaining competitive performance while offering a balance of interpretability and complexity.

By comparison, simpler models like logistic regression have shown lower performance. For example, Hasnain et al. (93) used logistic regression alongside more complex models like SVM and Random Forest, finding it lagged in predictive power. While easier to interpret, simpler models often struggle with multi-modal and high-dimensional data compared to more advanced algorithms. Another notable trend is the integration of traditional statistical models with ML techniques, such as in Jobczyk et al. (92), where a Cox proportional hazards model was combined with a deep neural network. This hybrid approach allows models to retain the interpretability of traditional methods while leveraging the predictive power of modern AI techniques, a promising direction for clinical adoption.

Building on our review, we highlight the transformative potential of integrating ML-based models into NMIBC management. Below, we outline key opportunities and challenges that must be addressed for successful clinical adoption.

4.1 Opportunities and transformative potential

- **Enhanced Accuracy and Precision:** One of the most compelling opportunities presented by ML in NMIBC management is the potential for improved diagnostic accuracy and precision. ML models can identify subtle patterns and complex relationships that may be

overlooked by human observers, reducing interobserver variability and mitigating the influence of individual clinician biases. This enhanced consistency can lead to better patient outcomes through more accurate predictions of recurrence and progression.

- **Automation Leading to Productivity and Cost Reduction:** ML algorithms can automate complex and time-consuming tasks such as image analysis, pattern recognition, and data integration. Automation increases productivity by enabling faster processing of large volumes of data and reduces healthcare costs through more efficient resource utilization. This allows clinicians to focus more on patient care rather than administrative or repetitive tasks.
- **Reduction of Interobserver Variability and Clinician Bias:** Diagnoses often rely on each clinician's personal experience and are subject to interobserver variability and bias. ML models can provide consistent and objective analyses, reducing variability between clinicians and minimizing the impact of subjective judgment.
- **Personalized Medicine and Transferable Knowledge:** ML models facilitate personalized treatment strategies by analyzing individual patient data to predict responses to specific therapies. Tailoring treatment plans to each patient's unique profile can improve efficacy and reduce unnecessary interventions. Additionally, the methodologies developed through ML in NMIBC are transferable to muscle-invasive bladder cancer (MIBC) and other malignancies, broadening the impact of this technology. ML tools can also serve as educational resources for training new professionals, enhancing learning through explainable models that illustrate decision-making processes.

4.2 Challenges and considerations

Despite the encouraging results presented in multiple studies, several significant hurdles exist before ML-based models can be successfully implemented in clinical settings. Below, we discuss key challenges identified in our review.

- **Limited Generalizability and Overfitting:** AI models often struggle with the heterogeneity of NMIBC tumors in stage, grade, and molecular subtypes. Small sample sizes in many studies lead to overfitted models that are not representative of the broader NMIBC population. High predictive accuracy reported by studies like Ajili et al. is limited by factors such as retrospective designs and single-institution datasets, which affect generalizability. To address this, multi-institutional data, diverse imaging techniques, and a wide range of NMIBC presentations should be included in training datasets.

- **Complexity of Implementation in Clinical Settings:** Implementing complex multimodal models in clinical practice requires specialized expertise, which may not always be available. Multimodal models need more complex data preprocessing and tuning, making scaling difficult across centers. Additionally, the lack of transparency in some models, particularly deep learning, hinders their adoption in clinical practice.
- **Lack of Interpretability and the “Black-Box” Nature of AI Models:** Clinicians need models that not only provide accurate predictions but also explain their reasoning to support confident, informed medical decisions. This challenge highlights a broader issue in the field of ML in healthcare: a significant disconnect between ML research and clinical requirements. While researchers often prioritize high accuracy, favoring complex models that enhance publication appeal, clinicians prioritize models that are interpretable, even if they sacrifice *some* accuracy. Bridging this gap requires a focus on developing models that balance performance with interpretability, better aligning ML innovations with the practical needs of real-world clinical applications.
- **Regulatory Challenges and Algorithmic Bias:** Predicting NMIBC recurrence faces challenges such as data access regulations, patient confidentiality, and ethical approvals, which can delay research and limit data availability. Additionally, ML algorithms may perpetuate biases, reducing prediction reliability for underrepresented groups. Addressing these issues requires implementing *federated learning* to train models across institutions without sharing raw data, using *differential privacy* to anonymize sensitive information, and ensuring NMIBC datasets reflect diverse patient demographics to improve accuracy and fairness.

4.3 Future directions

To fully leverage ML in NMIBC management, future efforts should prioritize developing multimodal models, improving interpretability, integrating ML into clinical workflows, and addressing ethical and regulatory barriers.

- **Multimodal Models:** Combining genomics, radiomics, and clinical data can enhance predictive power and generalizability. International collaboration is essential to assemble large, diverse datasets that capture variability in patient demographics and practices. Federated learning can enable training on multi-center data while preserving patient privacy, thus boosting model robustness.
- **Improving Interpretability:** To build clinician trust, models should incorporate explainable AI techniques, such as SHAP

values or attention mechanisms, making predictions transparent. User-friendly interfaces with visualization tools can further support clinician interaction with model outputs, enhancing usability in practice.

- **Integrating ML into Clinical Workflows:** Embedding ML tools within Electronic Health Records can offer real-time decision support. Adaptive decision-support systems that refine recommendations based on clinician feedback and evolving patient data will enhance decision-making by incorporating both clinical expertise and algorithmic insights.
- **Regulatory Sandbox for ML Testing:** Establishing a regulatory sandbox can accelerate validation by allowing controlled testing in clinical settings with flexible oversight. This framework supports iterative model adjustments and real-time data collection, speeding up the approval process. Engaging regulatory bodies to develop NMIBC-specific guidelines will facilitate smoother pathways to clinical adoption.

5 Conclusion

NMIBC, with its high recurrence rate of 70%-80% and substantial treatment costs, demands innovative predictive solutions. Our systematic review of 25 ML-based studies highlights the transformative potential of ML in NMIBC management. We explore various ML-based frameworks that utilize radiomics, histopathological markers, clinical data, genomics, and their combinations to predict NMIBC recurrence. Studies that integrated multiple data sources demonstrated remarkable accuracy, with NNs leading the charge. Our review encompasses the usage of ML-based models, acknowledges the potential for failures, and emphasizes the need for further intensive investigations to ensure their beneficial application. Unlike previous reviews that leave researchers struggling with complex technical details, our in-depth and nuanced evaluation of each study’s methodologies offers simplified, valuable insights into the intricacies of these ML algorithms, their clinical relevance, and practical applications. The future of NMIBC management is poised for innovation, as ML models have the potential to reduce both bias and interobserver variability. However, a significant hurdle remains - the scarcity of high-quality data. Developing robust and expansive datasets through collaboration is crucial for training models that can deliver real-world impact. While ML has shown immense potential in reshaping personalized medicine, its role in predicting NMIBC recurrence is not yet the gold standard.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Author contributions

SA: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. NS: Formal analysis, Investigation, Methodology, Supervision, Writing – review & editing. RH: Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Writing – review & editing. RS: Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing – review & editing. KA: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research received funding from EPSRC EP/X036006/1 to cover the publication costs.

References

1. *Bladder Cancer Statistics* (2015). Available online at: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer> (Accessed June 02, 2023).
2. Van Den Bosch S, Alfred Witjes J. Long-term cancer-specific survival in patients with high-risk, non-muscle-invasive bladder cancer and tumour progression: A systematic review. *Eur Urol.* (2011) 60:493–500. doi: 10.1016/j.eururo.2011.05.045
3. Babjuk M, Böhle A, Burger M, Capoun O, Cohen D, Compérat EM, et al. EAU guidelines on non-muscle-invasive urothelial carcinoma of the bladder: update 2016. *Eur Urol.* (2017) 71:447–61. doi: 10.1016/j.eururo.2016.05.041
4. Hall MC, Chang SS, Dalbagni G, Pruthi RS, Seigne JD, Skinner EC, et al. Guideline for the management of nonmuscle invasive bladder cancer (Stages ta, T1, and tis): 2007 update. *J Urol.* (2007) 178:2314–30. doi: 10.1016/j.juro.2007.09.003
5. Witjes JA, Compérat E, Cowan NC, De Santis M, Gakis G, Lebre T, et al. EAU guidelines on muscle-invasive and metastatic bladder cancer: summary of the 2013 guidelines. *Eur Urol.* (2014) 65:778–92. doi: 10.1016/j.eururo.2013.11.046
6. Mossanen M, Gore JL. The burden of bladder cancer care: Direct and indirect costs. *Curr Opin Urol.* (2014) 24:487–91. doi: 10.1097/MOU.0000000000000078
7. Leal J, Luengo-Fernandez R, Sullivan R, Witjes JA. Economic burden of bladder cancer across the european union. *Eur Urol.* (2016) 69:438–47. doi: 10.1016/j.eururo.2015.10.024
8. Tan WS, Teo CH, Chan D, Heinrich M, Feber A, Sarpong R, et al. Mixed-methods approach to exploring patients' perspectives on the acceptability of a urinary biomarker test in replacing cystoscopy for bladder cancer surveillance. *Bju Int.* (2019) 124:408–17. doi: 10.1111/bju.14690
9. Lo E, Nicolle LE, Coffin SE, Gould C, Maragakis LL, Meddings J, et al. Strategies to prevent catheter-associated urinary tract infections in acute care hospitals: 2014 update. *Infection Control Hosp Epidemiol.* (2014) 35:464–79. doi: 10.1086/675718
10. *NHS England National Cost Collection for the NHS* (2021). Available online at: <https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/> (Accessed November 20, 2023).
11. *NHS England 2020/21 National Cost Collection Data Publication* (2020). Available online at: <https://www.england.nhs.uk/publication/2020-21-national-cost-collection-data-publication/> (Accessed November 20, 2023).
12. *NHS England 2019/20 National Cost Collection Data Publication* (2019). Available online at: <https://www.england.nhs.uk/publication/2019-20-national-cost-collection-data-publication/> (Accessed November 20, 2023).
13. *NHS England 2018/19 National Cost Collection Data Publication* (2018). Available online at: <https://www.england.nhs.uk/publication/2018-19-national-cost-collection-data-publication/> (Accessed November 20, 2023).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

14. *NHS reference costs 2015 to 2016* (2015). Available online at: <https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016> (Accessed November 20, 2023).
15. *Inflation Calculator* (2023). Available online at: <https://www.bankofengland.co.uk/monetary-policy/inflation/inflation-calculator> (Accessed November 20, 2023).
16. Cox E, Saramago P, Kelly J, Porta N, Hall E, Tan WS, et al. The Impacts of Bladder Cancer on UK Healthcare Costs and Patients' Health-Related Quality of Life: Evidence from the BOXIT Trial: Bladder cancer cost and health-related impacts. *Clin Genitourinary Cancer.* (2020) 18:(2019). doi: 10.1016/j.clgc.2019.12.004
17. Kandori S, Kojima T, Nishiyama H. The updated points of TNM classification of urological cancers in the 8th edition of AJCC and UICC. *Japanese J Clin Oncol.* (2019) 49:421–5. doi: 10.1093/jjco/hyz017
18. Edge SB, Compton CC. The american joint committee on cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol.* (2010) 17:1471–4. doi: 10.1245/s10434-010-0985-4
19. Chang SS, Boorjian SA, Chou R, Clark PE, Daneshmand S, Konety BR, et al. Diagnosis and treatment of non-muscle invasive bladder cancer: AUA/SUO guideline. *J Urol.* (2016) 196:1021–9. doi: 10.1016/j.juro.2016.06.049
20. Hensley PJ, Panebianco V, Pietzak E, Kutikov A, Vikram R, Galsky MD, et al. Contemporary staging for muscle-invasive bladder cancer: accuracy and limitations. *Eur Urol Oncol.* (2022) 5:403–11. doi: 10.1016/j.euo.2022.04.008
21. Kamat AM, Hahn NM, Efsthathiou JA, Lerner SP, Malmström PU, Choi W, et al. Bladder cancer. *Lancet.* (2016) 388:2796–810. doi: 10.1016/S0140-6736(16)30512-8
22. Clark PE, Agarwal N, Biagioli MC, Eisenberger MA, Greenberg RE, Herr HW, et al. Clinical practice guidelines in oncology. *J Natl Compr Cancer Network.* (2013) 11:460–1. doi: 10.6004/jnccn.2013.0059
23. Richards KA, Smith ND, Steinberg GD. The importance of transurethral resection of bladder tumor in the management of nonmuscle invasive bladder cancer: A systematic review of novel technologies. *J Urol.* (2014) 191:1655–64. doi: 10.1016/j.juro.2014.01.087
24. Burger M, Catto JW, Dalbagni G, Grossman HB, Herr H, Karakiewicz P, et al. Epidemiology and risk factors of urothelial bladder cancer. *Eur Urol.* (2013) 63:234–41. doi: 10.1016/j.eururo.2012.07.033
25. Chen HI, Liou SH, Loh CH, Uang SN, Yu YC, Shih TS. Bladder cancer screening and monitoring of 4,4'-Methylenebis(2-chloroaniline) exposure among workers in Taiwan. *Urology.* (2005) 66:305–10. doi: 10.1016/j.urology.2005.02.031
26. Letášiová S, Medve'ová A, Šovčíková A, Dušinská M, Volkovová K, Mosoiu C, et al. Bladder cancer, a review of the environmental risk factors. *Environ Health.* (2012) 11:S11. doi: 10.1186/1476-069X-11-S1-S11

27. Cumberbatch MGK, Jubber I, Black PC, Esperto F, Figueroa JD, Kamat AM, et al. Epidemiology of bladder cancer: A systematic review and contemporary update of risk factors in 2018. *Eur Urol.* (2018) 74:784–95. doi: 10.1016/j.euro.2018.09.001
28. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer J Clin.* (2018) 68:394–424. doi: 10.3322/caac.21492
29. Farling KB. Bladder cancer: Risk factors, diagnosis, and management. *Nurse Practitioner.* (2017) 42:26. doi: 10.1097/01.NPR.0000512251.61454.5c
30. Audenet F, Attalla K, Sfakianos JP. The evolution of bladder cancer genomics: What have we learned and how can we use it? *Urologic Oncology: Semin Original Investigations.* (2018) 36:313–20. doi: 10.1016/j.urolonc.2018.02.017
31. EORTC risk tables: Predicting recurrence and progression in stage Ta T1 bladder cancer. *Evidencio* (2023). Available online at: <https://www.evidencio.com/models/show/1025> (Accessed November 20, 2023).
32. Seo KW, Kim BH, Park CH, Kim CI, Chang HS. The efficacy of the EORTC scoring system and risk tables for the prediction of recurrence and progression of non-muscle-invasive bladder cancer after intravesical bacillus calmette-guerin instillation. *Korean J Urol.* (2010) 51:165–70. doi: 10.4111/kju.2010.51.3.165
33. Predicting disease recurrence and progression - Uroweb (2023). Available online at: <https://uroweb.org/guidelines/non-muscle-invasive-bladder-cancer/chapter/predicting-disease-recurrence-and-progression> (Accessed November 19, 2023).
34. Vedder MM, Márquez M, de Bekker-Grob EW, Calle ML, Dyrskjot L, Kogevinas M, et al. Risk prediction scores for recurrence and progression of non-muscle invasive bladder cancer: an international validation in primary tumours. *PLoS One.* (2014) 9:e96849. doi: 10.1371/journal.pone.0096849
35. Krajewski W, Aumatell J, Subiela JD, Nowak Ł, Tukiendorf A, Moschini M, et al. Accuracy of the CUETO, EORTC 2016 and EAU 2021 scoring models and risk stratification tables to predict outcomes in high-grade non-muscle-invasive urothelial bladder cancer. *Urologic Oncology: Semin Original Investigations.* (2022) 40:491.e11–491.e19. doi: 10.1016/j.urolonc.2022.06.008
36. Fujii Y. Prediction models for progression of non-muscle-invasive bladder cancer: A review. *Int J Urol.* (2018) 25:212–8. doi: 10.1111/iju.13509
37. Xylinas E, Kent M, Kluth L, Pycha A, Comploj E, Svatek RS, et al. Accuracy of the EORTC risk tables and of the CUETO scoring model to predict outcomes in non-muscle-invasive urothelial carcinoma of the bladder. *Br J Cancer.* (2013) 109:1460–6. doi: 10.1038/bjc.2013.372
38. Born K, Levinson W, De Freitas L. Reducing harm from overuse of healthcare. *BMJ.* (2022) 379:o2787. doi: 10.1136/bmj.o2787
39. Morton B. Warning over cuts to NHS services without £10bn extra funding. *BBC News.* (2021).
40. Greenfield P. NHS wields the axe on 17 ‘unnecessary procedures’. *Guardian.* (2018).
41. Monteiro LL, Witjes JA, Agarwal PK, Anderson CB, Bivalacqua TJ, Bochner BH, et al. ICUD-SIU International Consultation on Bladder Cancer 2017: Management of non-muscle invasive bladder cancer. *World J Urol.* (2019) 37:51–60. doi: 10.1007/s00345-018-2438-9
42. Kim LHC, Patel MI. Transurethral resection of bladder tumour (TURBT). *Trans Andrology Urol.* (2020) 9:3056–72. doi: 10.21037/tau.2019.09.38
43. Zhu S, Yu W, Yang X, Wu C, Cheng F. Traditional classification and novel subtyping systems for bladder cancer. *Front Oncol.* (2020) 10:102. doi: 10.3389/fonc.2020.00102
44. Castaneda PR, Theodorescu D, Rosser CJ, Ahdo M. Identifying novel biomarkers associated with bladder cancer treatment outcomes. *Front Oncol.* (2023) 13. doi: 10.3389/fonc.2023.1114203
45. Zhang C, Hu J, Li H, Ma H, Othmane B, Ren W, et al. Emerging biomarkers for predicting bladder cancer lymph node metastasis. *Front Oncol.* (2021) 11. doi: 10.3389/fonc.2021.648968
46. Huang LK, Lin YC, Chuang HH, Chuang CK, Pang ST, Wu CT, et al. Body composition as a predictor of oncological outcome in patients with non-muscle-invasive bladder cancer receiving intravesical instillation after transurethral resection of bladder tumor. *Front Oncol.* (2023) 13:1180888. doi: 10.3389/fonc.2023.1180888
47. Xu X, Wang H, Du P, Zhang F, Li S, Zhang Z, et al. A predictive nomogram for individualized recurrence stratification of bladder cancer using multiparametric MRI and clinical risk factors. *J Magnetic Resonance Imaging.* (2019) 50:1893–904. doi: 10.1002/jmri.26749
48. Shkolyar E, Jia X, Chang TC, Trivedi D, Mach KE, Meng MQH, et al. Augmented bladder tumor detection using deep learning. *Eur Urol.* (2019) 76:714–8. doi: 10.1016/j.euro.2019.08.032
49. Tokuyama N, Saito A, Muraoka R, Matsubara S, Hashimoto T, Satake N, et al. Prediction of non-muscle invasive bladder cancer recurrence using machine learning of quantitative nuclear features. *Modern Pathol.* (2022) 35:533–8. doi: 10.1038/s41379-021-00955-y
50. Pantazopoulos D, Karakitsos P, Iokim-liosis A, Pouliakis A, Botsoli-stergiou E, Dimopoulos C. Back propagation neural network in the discrimination of benign from malignant lower urinary tract lesions. *J Urol.* (1998) 159:1619–23. doi: 10.1097/00005392-199805000-00057
51. Shao CH, Chen CL, Lin JY, Chen CJ, Fu SH, Chen YT, et al. Metabolite marker discovery for the detection of bladder cancer by comparative metabolomics. *Oncotarget.* (2017) 8:38802–10. doi: 10.18632/oncotarget.16393
52. Catto JWF, Linkens DA, Abbot MF, Chen M, Burton JL, Feeley KM, et al. Artificial intelligence in predicting bladder cancer outcome: A comparison of neuro-fuzzy modeling and artificial neural networks. *Clin Cancer Res.* (2003) 9:4172–7. doi: 10.1016/S1569-9056(03)80262-2
53. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* (2017) 14:749–62. doi: 10.1038/nrclinonc.2017.141
54. Zheng Q, Yang R, Ni X, Yang S, Xiong L, Yan D, et al. Accurate diagnosis and survival prediction of bladder cancer using deep learning on histological slides. *Cancers.* (2022) 14:5807. doi: 10.3390/cancers14235807
55. Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep.* (2018) 8:3395. doi: 10.1038/s41598-018-21758-3
56. Yuan Y, Shi Y, Li C, Kim J, Cai W, Han Z, et al. DeepGene: An advanced cancer type classifier based on deep learning and somatic point mutations. *BMC Bioinf.* (2016) 17:476. doi: 10.1186/s12859-016-1334-9
57. Mobadersany P, Yousefi S, Amgad M, Gutman DA, Barnholtz-Sloan JS, Velázquez Vega JE, et al. Predicting cancer outcomes from histology and genomics using convolutional neural networks. *Proc Natl Acad Sci.* (2018) 115:E2970–9. doi: 10.1073/pnas.1717139115
58. Cha KH, Hadjiiski LM, Samala RK, Chan HP, Cohan RH, Caoili EM, et al. Bladder cancer segmentation in CT for treatment response assessment: application of deep-learning convolution neural network—A pilot study. *Tomography.* (2016) 2:421–9. doi: 10.18383/j.tom.2016.00184
59. Jansen I, Lucas M, Bosschiet J, de Boer OJ, Meijer SL, van Leeuwen TG, et al. Automated detection and grading of non-muscle-invasive urothelial cell carcinoma of the bladder. *Am J Pathol.* (2020) 190:1483–90. doi: 10.1016/j.ajpath.2020.03.013
60. Wang H, Hu D, Yao H, Chen M, Li S, Chen H, et al. Radiomics analysis of multiparametric MRI for the preoperative evaluation of pathological grade in bladder cancer tumors. *Eur Radiol.* (2019) 29:6182–90. doi: 10.1007/s00330-019-06222-8
61. Mucaki EJ, Zhao JZL, Lizotte DJ, Rogan PK. Predicting responses to platinum chemotherapy agents with biochemically-inspired machine learning. *Signal Transduction Targeted Ther.* (2019) 4:1–12. doi: 10.1038/s41392-018-0034-5
62. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med.* (2009) 6:e1000100. doi: 10.1371/journal.pmed.1000100
63. Chen S, Jiang L, Zheng X, Shao J, Wang T, Zhang E, et al. Clinical use of machine learning-based pathomics signature for diagnosis and survival prediction of bladder cancer. *Cancer Sci.* (2021) 112:2905–14. doi: 10.1111/cas.14927
64. Drachneris J, Rasmussen A, Morkunas M, Fabijonavicius M, Cekauskas A, Jankevicius F, et al. CD8+ Cell Density Gradient across the Tumor Epithelium–Stromal Interface of Non-Muscle Invasive Papillary Urothelial Carcinoma Predicts Recurrence-Free Survival after BCG Immunotherapy. *Cancers.* (2023) 15:1205. doi: 10.3390/cancers15041205
65. Levy JJ, Chan N, Marotti JD, Rodrigues NJ, Ismail AAO, Kerr DA, et al. Examining longitudinal markers of bladder cancer recurrence through a semiautonomous machine learning system for quantifying specimen atypia from urine cytology. *Cancer Cytopathology.* (2023) 131:561–73. doi: 10.1002/cncy.22725
66. Wang GY, Zhu JF, Wang QC, Qin JX, Wang XL, Liu X, et al. Prediction of non-muscle invasive bladder cancer recurrence using deep learning of pathology image. *Sci Rep.* (2024) 14:18931. doi: 10.1038/s41598-024-66870-9
67. Huang H, Huang Y, Kaggie JD, Cai Q, Yang P, Wei J, et al. Multiparametric MRI-based deep learning radiomics model for assessing 5-year recurrence risk in non-muscle invasive bladder cancer. *J Magnetic Resonance Imaging.* (2024). doi: 10.1002/jmri.29574
68. Wang H, Zhang M, Miao J, Hou F, Chen Y, Huang Y, et al. Deep learning signature based on multiphase enhanced CT for bladder cancer recurrence prediction: A multi-center study. *eClinicalMedicine.* (2023) 66:102352. doi: 10.1016/j.eclinm.2023.102352
69. Lee J, Choo MS, Yoo S, Cho MC, Son H, Jeong H. Intravesical prostatic protrusion and prognosis of non-muscle invasive bladder cancer: analysis of long-term data over 5 years with machine-learning algorithms. *J Clin Med.* (2021) 10:4263. doi: 10.3390/jcm10184263
70. Zhao H, Chen Z, Fang Y, Su M, Xu Y, Wang Z, et al. Prediction of prognosis and recurrence of bladder cancer by ECM-related genes. *J Immunol Res.* (2022) 2022:1793005. doi: 10.1155/2022/1793005
71. Cai T, Nesi G, Canto MD, Mondaini N, Piazzini M, Bartoletti R. Prognostic role of loss of heterozygosity on chromosome 18 in patients with low-risk nonmuscle-invasive bladder cancer: results from a prospective study. *J Surg Res.* (2010) 161:89–94. doi: 10.1016/j.jss.2008.10.017
72. Sivic D, Taubert H, Breyer J, Eckstein M, Weyerer V, Keck B, et al. The prognostic value of FGFR3 expression in patients with T1 non-muscle invasive bladder cancer. *Cancer Manage Res.* (2021) 13:6567–65780. doi: 10.2147/CMAR.S318893

73. Bertz S, Otto W, Denzinger S, Wieland WF, Burger M, Stöhr R, et al. Combination of CK20 and ki-67 immunostaining analysis predicts recurrence, progression, and cancer-specific survival in pT1 urothelial bladder cancer. *Eur Urol.* (2014) 65:218–26. doi: 10.1016/j.eururo.2012.05.033
74. Wang Z, Que H, Suo C, Han Z, Tao J, Huang Z, et al. Evaluation of the NMP22 BladderChek test for detecting bladder cancer: A systematic review and meta-analysis. *Oncotarget.* (2017) 8:100648–56. doi: 10.18632/oncotarget.22065
75. Ponsky LE, Sharma S, Pandrangi L, Kedia S, Nelson D, Agarwal A, et al. Screening and monitoring for bladder cancer: refining the use of NMP22. *J Urol.* (2001) 166:75–8. doi: 10.1016/S0022-5347(05)66080-6
76. Urbanowicz RJ, Andrew AS, Karagas MR, Moore JH. Role of genetic heterogeneity and epistasis in bladder cancer susceptibility and outcome: A learning classifier system approach. *J Am Med Inf Assoc.* (2013) 20:603–12. doi: 10.1136/amiajnl-2012-001574
77. Wu X, Hildebrandt MAT, Chang DW. Genome-wide association studies of bladder cancer risk: A field synopsis of progress and potential applications. *Cancer Metastasis Rev.* (2009) 28:269–80. doi: 10.1007/s10555-009-9190-y
78. Menon DK, Rosand J. Finding a place for candidate gene studies in a genome-wide association study world. *JAMA Network Open.* (2021) 4:e2118594. doi: 10.1001/jamanetworkopen.2021.18594
79. Garcia-Closas M, Ye Y, Rothman N, Figueroa JD, Malats N, Dinney CP, et al. A genome-wide association study of bladder cancer identifies a new susceptibility locus within SLC14A1, a urea transporter gene on chromosome 18q12.3. *Hum Mol Genet.* (2011) 20:4282–9. doi: 10.1093/hmg/ddr342
80. Rafnar T, Sulem P, Thorleifsson G, Vermeulen SH, Helgason H, Saemundsdottir J, et al. Genome-wide association study yields variants at 20p12.2 that associate with urinary bladder cancer. *Hum Mol Genet.* (2014) 23:5545–57. doi: 10.1093/hmg/ddu264
81. Wang M, Li Z, Chu H, Lv Q, Ye D, Ding Q, et al. Genome-wide association study of bladder cancer in a chinese cohort reveals a new susceptibility locus at 5q12.3. *Cancer Res.* (2016) 76:3277–84. doi: 10.1158/0008-5472.CAN-15-2564
82. López de Maturana E, Picornell A, Masson-Lecomte A, Kogevinas M, Márquez M, Carrato A, et al. Prediction of non-muscle invasive bladder cancer outcomes assessed by innovative multimarker prognostic models. *BMC Cancer.* (2016) 16:351. doi: 10.1186/s12885-016-2361-7
83. Chang Q, Chen Y, Yin J, Wang T, Dai Y, Wu Z, et al. Comprehensive urinary proteome profiling analysis identifies diagnosis and relapse surveillance biomarkers for bladder cancer. *J Proteome Res.* (2024) 23:2241–52. doi: 10.1021/acs.jproteome.4c00199
84. Frantzi M, Van Kessel KE, Zwarthoff EC, Marquez M, Rava M, Malats N, et al. Development and validation of urine-based peptide biomarker panels for detecting bladder cancer in a multi-center study. *Clin Cancer Res.* (2016) 22:4077–86. doi: 10.1158/1078-0432.CCR-15-2715
85. Krochmal M, van Kessel KEM, Zwarthoff EC, Belczacka I, Pejchinovski M, Vlahou A, et al. Urinary peptide panel for prognostic assessment of bladder cancer relapse. *Sci Rep.* (2019) 9:7635. doi: 10.1038/s41598-019-44129-y
86. Zhan Y, Du L, Wang L, Jiang X, Zhang S, Li J, et al. Expression signatures of exosomal long non-coding RNAs in urine serve as novel non-invasive biomarkers for diagnosis and recurrence prediction of bladder cancer. *Mol Cancer.* (2018) 17:142. doi: 10.1186/s12943-018-0893-y
87. Gogalic S, Sauer U, Doppler S, Heinzel A, Perco P, Lukas A, et al. Validation of a protein panel for the noninvasive detection of recurrent non-muscle invasive bladder cancer. *Biomarkers.* (2017) 22:674–81. doi: 10.1080/1354750X.2016.1276628
88. Ajili F. Prognostic value of artificial neural network in predicting bladder cancer recurrence after BCG immunotherapy. *J Cytology Histol.* (2014) 5:1–5. doi: 10.4172/2157-7099.1000226
89. Zhang X, Ma L. Predictive value of the total bilirubin and CA50 screened based on machine learning for recurrence of bladder cancer patients. *Cancer Manage Res.* (2024) 16:537–46. doi: 10.2147/CMAR.S457269
90. Schwarz L, Sobania D, Rothlauf F. On relevant features for the recurrence prediction of urothelial carcinoma of the bladder. *Int J Med Inf.* (2024) 186:105414. doi: 10.1016/j.ijmedinf.2024.105414
91. Lucas M, Jansen I, van Leeuwen TG, Oddens JR, de Bruin DM, Marquering HA. Deep learning-based recurrence prediction in patients with non-muscle-invasive bladder cancer. *Eur Urol Focus.* (2022) 8:165–72. doi: 10.1016/j.euf.2020.12.008
92. Jobczyk M, Stawiski K, Kaszkowiak M, Rajwa P, Różański W, Soria F, et al. Deep learning-based recalibration of the CUETO and EORTC prediction tools for recurrence and progression of non-muscle-invasive bladder cancer. *Eur Urol Oncol.* (2022) 5:109–12. doi: 10.1016/j.euo.2021.05.006
93. Hasnain Z, Mason J, Gill K, Miranda G, Gill IS, Kuhn P, et al. Machine learning models for predicting post-cystectomy recurrence and survival in bladder cancer patients. *PLoS One.* (2019) 14:e0210976. doi: 10.1371/journal.pone.0210976