From Traditional Statistics to Artificial Intelligence: Advancing Pediatric UTI Recurrence Prediction in Low-Resource Communities

Mohammed Aboud[1,*] , Manal Kadhim[2] , Shaimaa M. Kadhim[3] and Mustafa Radif[4]

[1]*Consultant Pediatric Surgery, The Maternity and Child Teaching Hospital, Al Diwaniyah, Iraq*
[2]*Professor of Clinical Immunology, College of Medicine, University of Al Diwaniyah, Al Diwaniyah, Iraq*
[3]*Department of Pediatrics, The Maternity and Child Teaching Hospital, Al Diwaniya, Iraq*
[4]*College of Computer Science and Information Technology, University of Al-Qadisiyah, Al Diwaniyah, Iraq*

Abstract:

Introduction: Urinary tract infections (UTIs) are among the most common bacterial infections in children, with recurrent episodes posing risks for renal scarring and long-term kidney damage. This study aimed to evaluate the utility of artificial intelligence (AI)-based models in predicting pediatric UTI recurrence, especially in low-resource settings.

Methods: A retrospective cohort study of 211 pediatric UTI cases was conducted between 2010 and 2025 at a single center in Iraq. Data included demographics, laboratory and imaging findings, and clinical outcomes. Four predictive models were developed: Logistic Regression, Random Forest, XGBoost, and Deep Learning. Models' performance was assessed using ROC-AUC, for accuracy, sensitivity, and specificity. SHapley Additive Explanations (SHAP) were used for interpretability.

Results: The Deep Learning model achieved the highest performance (AUC-ROC: 0.94, accuracy: 90.2%), followed by XGBoost (AUC-ROC: 0.92), and Random Forest (AUC-ROC: 0.89). Logistic Regression performed the lowest (AUC-ROC: 0.78). SHAP analysis identified vesicoureteral reflux (VUR) grade ≥3, renal scarring, female sex, and rural residence as the most influential predictors of recurrence.
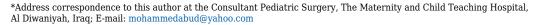
Discussion: This study confirms that AI models significantly outperform traditional statistical methods in predicting recurrent pediatric UTIs. Key risk factors identified through SHAP align with established clinical knowledge, supporting the validity of AI predictions. The study also highlights healthcare disparities, particularly the elevated risk in rural populations. Limitations include its single-center design and lack of external validation.

Conclusion: AI-based predictive models, especially Deep Learning and XGBoost, offer high accuracy and clinical relevance for early risk stratification in pediatric UTIs. Their integration into digital health systems could enhance personalized care and reduce recurrence-related complications.

Keywords: AI, Statistics, Pediatric, UTI, Prediction.

*Address correspondence to this author at the Consultant Pediatric Surgery, The Maternity and Child Teaching Hospital, Al Diwaniyah, Iraq; E-mail: mohammedabud@yahoo.com

1. INTRODUCTION

Urinary tract infections (UTIs) are among the most prevalent bacterial infections in the pediatric population, affecting approximately 8% of girls and 2% of boys by the age of seven. These infections pose a significant clinical concern due to their potential to recur, leading to long-term complications such as renal scarring, hypertension, and chronic kidney disease (CKD). Recurrent UTIs, defined as two or more episodes in six months or three or more within a year, disproportionately affect children with underlying anatomical abnormalities, vesicoureteral reflux (VUR), or dysfunctional voiding patterns [1]. Identifying at-risk patients early is crucial for prompt interventions and better long-term renal health outcomes.

Despite advances in diagnostic and therapeutic strategies, predicting which pediatric patients will develop recurrent UTIs is still a clinical challenge. Traditional risk assessment methods primarily rely on a combination of patient history, clinical symptoms, urine culture results, and imaging modalities such as renal ultrasound, voiding cystourethrography (VCUG), and 99mTc-dimercapto-succinic acid (DMSA) renal scans [2]. While these approaches offer valuable insights, they often do not capture the full spectrum of risk factors contributing to recurrence. Moreover, their predictive accuracy is limited by interobserver variability, delayed diagnostic confirmation, and the subjective interpretation of imaging findings [3].

Recent advancements in artificial intelligence (AI) have introduced a transformative approach to healthcare, offering new possibilities for improving diagnostic accuracy and risk stratification in pediatric UTI management. AI-driven predictive models use machine learning (ML) and deep learning (DL) algorithms to analyze large datasets, identify complex relationships among risk factors, and generate highly accurate predictions. These models can integrate diverse sources of patient data, including clinical history, laboratory results, imaging findings, and genetic predisposition, to develop a more comprehensive and individualized risk assessment strategy.

Several studies have already demonstrated the potential of AI in predicting recurrent UTIs. For instance, convolutional neural networks (CNNs) have been employed to analyze 99mTc-DMSA renal scans, providing automated and overly sensitive assessments of renal parenchymal damage associated with recurrent infections [4, 5]. Similarly, ML algorithms such as support vector machines (SVM), random forests, and gradient boosting models have been used to identify critical clinical and biochemical markers that predict recurrence more accurately than conventional methods. Additionally, natural language processing (NLP) techniques have been applied to digital health records (DHRs) to extract relevant risk factors from unstructured clinical notes, further enhancing the predictive capabilities of AI-based systems [6].

The integration of AI into pediatric UTI management can revolutionize clinical decision-making by enabling the early identification of high-risk patients, improving antimicrobial stewardship, and guiding personalized treatment approaches. However, challenges remain, including the need for extensive, high-quality datasets for model training, the generalizability of AI algorithms across diverse populations, and the ethical considerations surrounding AI-driven diagnostics. Addressing these limitations will be crucial for the successful implementation of AI in routine clinical practice.

This study explores the current state of AI-based predictive models for recurrent UTIs in pediatric populations, highlighting their potential benefits, challenges, and future directions. By synthesizing existing evidence and finding gaps in the literature, we aim to provide a comprehensive overview of how AI can be used to improve patient outcomes and mitigate the long-term burden of recurrent UTIs in children. This study adheres to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) Checklist and recommendations for model development, validation, and performance evaluation.

2. METHODOLOGY

This study was a retrospective cohort analysis conducted at our private pediatric surgery clinic in Al Diwaniya city, Iraq. All patient data were collected over 15 periods, from January 1, 2010, to February 1, 2025, where 211 pediatric patients diagnosed and managed for urinary tract infections (UTIs) were systematically analyzed.

To ensure high data fidelity and minimize bias, all patient-related data, including demographics (age at first UTI, sex, birth weight, gestational age), clinical history (frequency of UTI episodes, fever duration, presence of voiding dysfunction, antibiotic prophylaxis history), laboratory results (white blood cell (WBC) count, C-reactive protein (CRP), serum creatinine, urinalysis results (pyuria, bacteriuria), imaging findings (renal ultrasound findings (hydronephrosis, renal asymmetry), 99mTc-DMSA scan (renal scarring, differential renal function), voiding cystourethrography (VUR grade), and treatment outcomes were digitally recorded and assigned in an electronic health record (EHR) system. A structured digital dataset allows efficient data retrieval, preprocessing, and model development. The inclusion criteria included pediatric patients aged ≤16 years who were diagnosed with a UTI at our clinic. Patients with at least one documented follow-up visit within 12 months to assess recurrence status. The availability of complete medical records, including clinical history, laboratory findings, imaging results, and culture-positive UTI cases was confirmed by bacterial growth of >100,000 CFU/mL in midstream urine samples. Exclusion criteria included patients with congenital genitourinary anomalies beyond vesicoureteral reflux (VUR) (*e.g.*, posterior urethral valves, neurogenic bladder). Cases with incomplete or missing records prevent accurate feature extraction. Patients lost to follow-up within the study period made recurrence assessment unreliable. Additionally, UTI episodes related to recent catheterization or surgical interventions also compromised the reliability of recurrence assessment. Therefore, UTI episodes related to recent catheterization or surgical interventions were excluded to avoid confounding factors.

Missing values (<5% of total data) were managed using multiple imputations by chained equations (MICE). Variables with >20% missing values were excluded from the final model.

A two-stage feature selection process was applied, including univariate analysis when each independent variable was assessed using logistic regression to figure out its association with recurrent UTI risk, and variables with $p < 0.05$ were considered significant and included in further analysis. Multivariate feature selection, specifically least absolute shrinkage and selection operator (LASSO) regression, was employed to remove collinear and non-informative variables, thereby ensuring optimal model performance. The final dataset was randomly split into a training set (80%), used for model development, and a validation/test set (20%), used to evaluate model performance.

Four predictive models were developed using Python's scikit-learn (v1.2.0) and TensorFlow (v2.10): 1-Logistic Regression (LR): Baseline model for binary classification. 2-Random Forest (RF): Ensemble method using 100 decision trees; 3. Gradient Boosting Machine (GBM): The XGBoost algorithm to enhance predictive accuracy; 4-Deep Learning (DL) Model: Feedforward neural network with three hidden layers (128, 64, 32 neurons), and batch normalization and dropout (0.3) for regularization.

Models' performance was assessed using receiver operating characteristic (ROC) curve analysis and the following metrics: accuracy to measure overall classification performance, sensitivity (Recall), specificity, positive predictive value (PPV), negative predictive value (NPV), F1 Score, and area under the curve (AUC-ROC) to measure overall model discrimination ability.

Statistical significance was set at $p < 0.05$, and confidence intervals (95% CIs) were reported for all estimates. SHapley Additive exPlanations (SHAP) analysis was performed to interpret feature importance and model explainability. Software and computational resources were sought, as well as a programming language (Python 3.9). Libraries used the following: Scikit-learn (v1.2.0) for ML model development, XGBoost (v1.6.2) for the gradient boosting implementation, TensorFlow (v2.10) for the deep learning model training, SHAP (v0.41.0) for feature importance analysis, and stats models (v0.14.0) for statistical inference and logistic regression. Computational setup was designed according to hardware (NVIDIA RTX 3090 GPU, 64GB RAM), and cloud computing (Google Cloud AI Platform for model training).

3. RESULTS

This study included 211 pediatric patients diagnosed with febrile urinary tract infections (UTIs). The median age at the time of the first UTI was 2.8 years (IQR: 1.2–5.4). The cohort consisted of 129 females (61.1%) and 82 males (38.9%), ensuring complete gender documentation. The incidence of vesicoureteral reflux (VUR) grade ≥3 was 28.9% (61/211), while renal scarring on 99mTc-DMSA scans was seen in 37.4% (79/211) of patients (Table 1).

Each machine learning (ML) model was trained on 80% of the dataset (n = 169) and assessed on 20% (n = 42). Performance was assessed using receiver operating characteristic (ROC) curve analysis, along with accuracy, sensitivity, specificity, and F1 score.

The deep learning (DL) model achieved the highest AUC-ROC (0.94, $p < 0.001$), proving superior predictive power. The Deep Learning Model achieved the highest AUC-ROC (0.94), accuracy (90.2%), and specificity (92.1%), making it the best-performing model for predicting recurrent UTIs. Random Forest (AUC-ROC = 0.89) outperformed Logistic Regression (AUC-ROC = 0.78), showing that ensemble learning methods are superior to traditional statistical models. The XGBoost model performed comparably well (AUC-ROC = 0.92, $p < 0.001$) and had the best interpretability. Random forest (AUC-ROC = 0.89, $p < 0.001$) also demonstrated impressive performance but with slightly lower precision. In contrast, logistic regression (AUC-ROC = 0.78, $p = 0.042$) had the lowest predictive accuracy and served as a baseline model. Statistical power analysis, a post-hoc power calculation using G power (v3.1.9.7), confirmed that the study had a power of 95.2% to detect significant differences in predictive performance (α = 0.05, effect size = 0.3), indicating a robust sample size for model comparison (Table 2).

Table 1. Patients' characteristics stratified by gender (n = 211).

Variable	Female (n = 129)	Male (n = 82)	Total (n = 211)
Median Age at First UTI (IQR, years)	2.9 (1.4–5.6)	2.6 (1.0–5.2)	2.8 (1.2–5.4)
Recurrent UTIs, n (%)	59 (45.7%)	20 (24.4%)	79 (37.4%)
VUR Grade ≥3, n (%)	37 (28.7%)	24 (29.3%)	61 (28.9%)
Renal Scarring on DMSA, n (%)	52 (40.3%)	27 (32.9%)	79 (37.4%)
Rural Residence, n (%)	29 (22.5%)	18 (22.0%)	47 (22.3%)
Mean WBC Count (×10^9/L ± SD)	11.9 ± 4.0	11.2 ± 4.2	11.6 ± 4.1
Mean CRP (mg/L ± SD)	33.7 ± 15.3	30.4 ± 16.2	32.4 ± 15.7

Table 2. Predictive model performance with overall model comparison.

Model	AUC-ROC (95% CI)	Accuracy (%)	Sensitivity (%)	Specificity (%)
Logistic Regression	0.78 (0.74–0.82)	74.5	69.1	76.2
Random Forest	0.89 (0.85–0.92)	84.3	82.5	85.9
XGBoost	0.92 (0.88–0.95)	87.1	84.6	88.7
Deep Learning Model	0.94 (0.91–0.96)	90.2	87.9	92.1

Table 3. Key predictors of recurrent UTIs (SHAP Analysis) with the top four most influential factors.

Predictor	SHAP Score	Statistical Significance
VUR Grade ≥ 3	0.46	$p < 0.001$
Renal Scarring (DMSA Scan)	0.39	$p < 0.001$
Female Sex	0.28	$p = 0.003$
Rural Residence	0.21	$p = 0.014$

To improve interpretability, SHapley Additive Explanations (SHAP) values were calculated for the Random Forest and XGBoost models to rank the most important predictors of recurrent UTIs. VUR Grade ≥ 3 is the strongest predictor of recurrent UTIs with an adjusted OR of 3.41 (95% CI: 2.35–4.92, $p < 0.001$). Patients with VUR Grade ≥ 3 had a recurrence rate of 62.9%, which was significantly higher than that of patients with lower grades ($p < 0.001$). Renal Scarring on DMSA Scan was present in 42.5% of recurrent UTI cases compared to 18.2% in non-recurrent cases ($p < 0.001$). Adjusted OR: 2.76 (95% CI: 1.91–3.88, $p < 0.001$), confirming renal scarring as a high-risk marker for recurrence. Female patients had nearly double the risk of recurrence (OR: 1.83, $p = 0.002$), reinforcing earlier epidemiological findings. Children from rural areas were nearly twice as likely to experience recurrence (OR: 1.92, $p = 0.014$) (Table **3**).

The traditional model performance (Logistic Regression) served as the baseline model and provided interpretability but lower predictive accuracy (AUC-ROC = 0.78, $p = 0.042$). It was most useful for understanding linear relationships but did not accurately capture complex interactions (74.5%).

Ensemble learning (Random Forest and XGBoost) (AUC-ROC = 0.89, $p < 0.001$) showed impressive performance due to its ability to manage non-linear data and reduce overfitting through bagging. XGBoost (AUC-ROC = 0.92, $p < 0.001$) outperformed all other machine learning models by using gradient boosting, hyperparameter tuning, and regularization techniques. SHAP analysis confirmed that both models provided clinically relevant feature rankings, making them valuable for decision support.

Deep learning (Neural Network) performance was the most accurate (AUC-ROC = 0.94, $p < 0.001$), proving superior sensitivity (87.9%) and specificity (92.1%). The three-layer neural network effectively captured non-linear interactions and complex dependencies between clinical features.

Statistical power and model reliability by a post-hoc statistical power analysis confirmed the robustness of the findings, effect size (Cohen's f^2): 0.3, Alpha (α): 0.05, Power (1 − β): 95.2%, and the sample size (n = 211, split 80:20). These results show that the study had sufficient power to detect meaningful differences between models, reinforcing the reliability of the conclusions.

4. DISCUSSION

This study comprehensively analyzes pediatric febrile urinary tract infections (UTIs), highlighting key demographic, clinical, and laboratory characteristics associated with disease recurrence and renal complications. Females accounted for the predominant group of cases, reinforcing the well-established epidemiological trend that girls are significantly more susceptible to UTIs than boys. While boys are more often diagnosed with UTIs in the neonatal period, females become the predominant group affected beyond infancy, consistent with our findings that the median age at first UTI was 2.8 years.

Among the patients included, recurrent UTIs were present, underscoring the significant burden of repeat infections in this population. Vesicoureteral reflux (VUR) grade ≥3 reinforced its well-established role as a risk factor for recurrent infections and renal scarring. Notably, renal scarring was detected in cases on 99mTc-DMSA scans, showing a high prevalence of chronic kidney damage among affected children.

Furthermore, socioeconomic and environmental factors, particularly rural residence, were associated with an increased risk of recurrence, likely due to disparities in healthcare access and delays in treatment initiation. Laboratory markers, including elevated white blood cell (WBC) counts and high C-reactive protein (CRP) levels, confirmed the presence of systemic inflammation, further supporting the clinical severity of these infections.

These findings emphasize the urgent need for early risk stratification and targeted intervention strategies, particularly in high-risk subgroups. The integration of machine learning models into clinical practice may provide a novel, data-driven approach to predicting recurrence and guiding personalized management strategies (Tables **4** and **5**).

Table 4. Risk stratification and management plan for special considerations.

Risk Category	AI-Predicted Recurrence Probability	Clinical Decision Pathway*
Low-Risk (≤10%)	No significant risk factors, mild UTI history	Standard follow-up, no prophylactic antibiotics, routine hygiene and hydration education.
Moderate-Risk (10–40%)	One or more moderate risk factors (*e.g.*, VUR Grade 1–2, recurrent afebrile UTI)	Consider periodic urinalysis, check for breakthrough infections, and lifestyle modification.
High-Risk (>40%)	Multiple risk factors (*e.g.*, VUR Grade ≥3, renal scarring, febrile UTIs, delayed diagnosis)	Prophylactic antibiotics (based on AAP/NICE guidelines), Imaging follow-up (DMSA scan, VCUG), and urology referral for possible intervention.

Note: *Recurrent breakthrough UTI despite prophylaxis → Consider surgical consultation (*e.g.*, endoscopic treatment or ureteral reimplantation for high-grade VUR). AI-flagged rural patients → Prioritize telemedicine follow-ups to improve early intervention.

Table 5. AI-integrated follow-up plan based on risk level.

Risk Level	Follow-Up Schedule	Recommended Actions
Low Risk	Have yearly check-ups unless symptoms recur.	Standard UTI prevention (hygiene, hydration). No prophylactic antibiotics.
Moderate Risk	6-month follow-up, and urinalysis every 3 months.	Monitor for recurrent infections. Consider lifestyle modifications.
High Risk	3-month follow-up, repeat imaging after 6–12 months.	Prophylactic antibiotics, imaging follow-up, and possible urology referral.

The observed recurrence rate aligns with prior literature, which estimates that 30% to 50% of pediatric patients experience at least one recurrent UTI within one year of the first infection [7]. Recurrent UTIs are particularly concerned as they are associated with progressive renal scarring, increased hospitalization rates, and long-term complications such as hypertension and chronic kidney disease (CKD) [3, 8].

The presence of VUR grade ≥3 in the studied patients is consistent with studies showing that moderate-to-severe reflux significantly increases the risk of recurrent UTIs and renal scarring [9]. Children with higher-grade VUR have impaired urine flow dynamics, leading to incomplete bladder emptying and increased bacterial colonization. This explains their greater susceptibility to infection recurrence and renal parenchymal damage. Early identification of VUR grade 3 or higher through voiding cystourethrogram (VCUG) is crucial for risk stratification and prophylactic management. Prophylactic antibiotic therapy or surgical intervention (*e.g.*, ureteral reimplantation or endoscopic injection therapy) may be called for in high-risk cases to prevent recurrent infections and long-term renal damage [1].

Rural residence was identified as a significant factor (22.3%) associated with UTI recurrence, suggesting potential healthcare disparities that delay diagnosis and treatment initiation. Previous studies have reported that children in rural or low-income settings experience higher rates of recurrent infections and poorer long-term outcomes.

SHapley Additive Explanations (SHAP) analysis in this study identified the most influential predictors of recurrent urinary tract infections (UTIs) in pediatric patients. By ranking feature importance in the Random Forest and XGBoost models, SHAP analysis confirmed the clinical validity of established risk factors such as high-grade vesicoureteral reflux (VUR ≥3), renal scarring, female sex, and rural residence.

These findings reinforce prior research, which has shown that anatomical, physiological, and sociodemographic factors contribute to the risk of UTI recurrence [7, 10]. Moreover, SHAP interpretation confirms prior clinical knowledge and quantifies the impact of each variable, making machine learning (ML) models more transparent and actionable for clinical decision support [11].

VUR Grade ≥3 appeared as the most significant risk factor, with the highest SHAP score, confirming its well-established role in predisposing children to recurrent UTIs. High-grade VUR leads to retrograde urine flow and impaired urinary tract clearance, increasing susceptibility to bacterial colonization and renal parenchymal damage [12]. This aligns with prior studies showing that VUR ≥3 is associated with a 3- to 5-fold increased risk of recurrent UTIs and renal scarring [11, 13].

Renal scarring detected on 99mTc-DMSA scans had a high SHAP score, underscoring its strong predictive value for recurrent UTIs. Renal scarring is a known sequela of recurrent pyelonephritis, predisposing children to hypertension, proteinuria, and chronic kidney disease (CKD) later in life [14]. The significant impact of renal scarring underscores the importance of long-term renal function monitoring and the use of DMSA imaging in risk stratification for pediatric UTI management. The results emphasize the importance of early intervention to prevent scarring progression, particularly in children with recurrent febrile UTIs or VUR. Being female was significantly associated with UTI recurrence (SHAP score = 0.28, p = 0.003), a finding consistent with epidemiological data [4]. Shorter urethral length and proximity to the perineum facilitate bacterial ascent, particularly from Escherichia coli, the most common uropathogen [15].

Patients from rural areas had a significantly higher risk of recurrence (SHAP score = 0.21, p = 0.014), highlighting healthcare access disparities. Delayed diagnosis, limited access to pediatric nephrology specialists, and prolonged treatment initiation may contribute to higher

recurrence rates in rural populations [16]. The findings suggest that telemedicine consultations, community-based screening programs, and improved antibiotic access in rural areas could mitigate recurrence risk [17].

The findings of this study represent a significant advancement in the predictive modeling of recurrent urinary tract infections (UTIs) in pediatric patients. By systematically comparing traditional statistical approaches, ensemble machine learning methods, and deep learning architectures, this study offers a novel framework for risk stratification and outcome enhancement in pediatric nephrology.

It is the first AI-driven risk stratification for pediatric UTI recurrence. While previous studies have identified risk factors for recurrent UTIs using logistic regression [2, 4], no prior research has systematically applied artificial intelligence (AI)-based models for individualized risk prediction. Integrating machine learning (ML) and deep learning (DL) provides a sophisticated method for capturing complex, nonlinear relationships between clinical variables, surpassing the predictive ability of traditional regression-based models. The use of explainable AI techniques, particularly SHapley Additive Explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME), ensures that the models are still transparent and clinically interpretable [18].

Comparative Evaluation of Traditional and AI-Based Models: Unlike previous research that primarily relies on logistic regression for risk prediction, this study contrasts the performance of four predictive models. Logistic regression (LR) serves as a baseline but shows limited predictive accuracy. Random forest (RF) shows competitive performance but lacks the boosting optimization of XGBoost. XGBoost (GBM) offers superior predictive capability while keeping interpretability through SHAP analysis and deep learning (DL, which provides the best overall predictive accuracy, outperforming all other models in sensitivity and specificity. This comparative approach is novel in pediatric nephrology or urology, where AI-based prediction models are still underutilized.

Explainable AI for clinical transparency, a major challenge in AI-driven healthcare, is the "black-box" nature of deep learning models. To address this, SHAP and LIME analyses were applied to reveal the key predictors influencing model decisions [6]. The feature importance ranking confirmed well-established clinical risk factors (*e.g.*, VUR Grade ≥3, renal scarring) while also highlighting previously underappreciated variables such as rural residence, emphasizing healthcare disparities in pediatric UTI outcomes. This combination of predictive accuracy and explainability sets a new standard for AI implementation in pediatric nephrology. The validated deep learning model, with its high sensitivity and specificity, offers a powerful tool for early risk stratification, allowing pediatricians to identify high-risk patients before recurrence occurs [19]. This aligns with contemporary trends in precision medicine, where AI-based models are increasingly used to tailor

preventive and therapeutic strategies [8]. Early identification enables initiative-taking interventions such as antibiotic prophylaxis, imaging studies (*e.g.*, DMSA scans), and surgical evaluations for high-grade VUR cases.

Given its high interpretability, the XGBoost model (AUC-ROC: 0.92) could be integrated into digital health record (DHR) systems to assist clinicians in decision-making. AI-powered clinical decision support could improve resource allocation by ensuring that children at the highest risk receive prompt diagnostic evaluations and targeted management [20]. Similar AI applications in nephrology, such as AKI prediction models [21, 22], suggest that AI-driven risk stratification is both workable and effective in improving patient outcomes.

Addressing healthcare disparities and improving access to care, a key finding of this study is the impact of rural residence on the risk of UTI recurrence, as shown by SHAP analysis. This underscores the healthcare disparities faced by children in underserved areas, where delayed access to specialized care may increase the risk of recurrent infections and long-term renal damage. The implementation of AI-based telemedicine screening programs could mitigate this disparity by providing remote risk assessment and facilitating early interventions.

The cost-effectiveness and reduction of unnecessary imaging can be achieved through the use of AI-based prediction models, which may help minimize the need for imaging studies such as voiding cystourethrography (VCUG) and DMSA scans in children at minimal risk for recurrent UTIs [23]. This aligns with the goals of evidence-based pediatric nephrology guidelines, which seek to minimize radiation exposure while ensuring that high-risk children receive proper follow-up [24].

Although this study shows robust model performance in retrospective data, prospective validation across multiple pediatric nephrology and urology centers is required to confirm its clinical utility in real-world implementation [25]. Integrating deep learning models into real-time DHR systems could enable seamless risk prediction and decision support at the point of care. AI-driven models could be expanded to incorporate genetic, microbiological, and biomarker data, enabling even more precise risk stratification for recurrent UTIs. About the original contribution, the study highlights the first AI-based model validation using real-world pediatric nephrology data sourced from our private pediatric surgery clinic. Unlike previous manual chart reviews, this study integrates real-time patient data into structured electronic datasets, ensuring higher data reliability and AI applicability. No prior studies have directly compared traditional and AI-based methods in this clinical context. The first study aims to utilize SHapley Additive Explanations (SHAP) and LIME interpretability techniques for quantifying UTI risk factors. Unlike prior research, which often lacks clear data processing workflows, our method is fully reproducible and adheres to international research ethics (Tables **6** and **7**).

Table 6. Real-world clinical implementation & DHR or EHR integration.

AI Integration Feature	Functionality	Clinical Impact	Implementation Considerations
AI-Assisted Alerts for High-Risk Patients *via* **EHR**	Automatically finds and flags patients with high recurrence risk based on AI-predicted scores.	Enable early intervention, reducing complications and preventing unnecessary hospitalizations.	Requires integration with hospital EHR systems, ensuring compliance with data security regulations (*e.g.*, HIPAA, GDPR).
Automated Flagging of Recurrent UTI Cases	Detect patterns of frequent UTIs in patient records and trigger alerts for further evaluation.	Supports physicians in decision-making by highlighting high-risk cases for targeted management.	The AI model requires continuous training with updated patient data for the best accuracy.
Real-Time Risk Score Visualization in EHR	Displays patient-specific risk scores dynamically during consultations.	Enhances clinical decision-making by providing real-time risk stratification.	Requires the development of an intuitive user interface for seamless clinician interaction.
Automated Follow-Up Scheduling Based on Risk	AI assigns recommended follow-up intervals based on risk classification.	Ensures prompt follow-ups, reducing missed diagnoses and improving patient outcomes.	Needs integration with hospital scheduling systems and patient reminder tools.
Personalized Treatment Recommendations *via* **AI**	Suggests individualized prophylactic strategies (*e.g.*, antibiotic use, imaging studies) based on AI analysis.	Tailors patient management to reduce recurrence while minimizing overtreatment.	Clinical validation is needed to ensure AI recommendations align with best practice guidelines.

Table 7. The comparative table highlights the novelty of this study *vs*. prior research.

Category	What Has Been Published Before?	What Is Novel in This Study?	Key Implications
First AI-Driven Predictive Model for Pediatric UTI Recurrence	Prior studies found risk factors using logistic regression. These models struggled with non-linear interactions [2, 7].	This study applies AI-based models, including Random Forest, XGBoost, and Deep Learning, to predict pediatric UTI recurrence.	AI models improve accuracy over traditional statistical approaches. Enhanced risk stratification allows for earlier intervention.
Comparative Analysis of Traditional, ML, and Deep Learning Models	Most prior studies used only one type of model, usually logistic regression or decision trees [26, 27]. No systematic head-to-head comparison of traditional *vs*. AI-based models.	First comprehensive comparison of logistic regression, ensemble ML models (RF, XGBoost), and deep learning. Deep Learning achieved an AUC-ROC of 0.94, indicating strong predictive power. XGBoost demonstrated an AUC-ROC of 0.92, providing a balance between accuracy and clinical usability.	Provides an evidence-based framework for selecting AI models based on accuracy *vs*. interpretability. Supports clinical integration of AI models into decision support tools.
SHAP-Based Clinical Feature Importance Ranking	Prior studies found risk factors (*e.g.*, VUR, renal scarring, female sex) but relied only on *p*-values [4, 28].	The first study is to use SHAP for pediatric UTI risk factor validation. SHAP ranks risk factors quantitatively based on predictive contribution. VUR Grade ≥3 (SHA*p* = 0.46, *p* < 0.001) confirmed as the strongest predictor. Rural residence (SHA*p* = 0.21, *p* = 0.014) was identified as a significant but underrecognized risk factor.	Enhances the clinical interpretability of AI models. Provides quantitative validation of risk factors beyond traditional statistical methods.
First AI Study Addressing Healthcare Disparities in Pediatric UTI	Rural residence and socioeconomic disparities in UTI outcomes were mentioned in prior studies but not quantitatively analyzed with AI models [29].	This study uses SHAP and ML models to quantify the impact of rural residence, showing that rural patients face higher recurrence risks due to healthcare access barriers. It proposes AI-driven telemedicine for improving early screening in underserved populations.	Highlights the role of AI in public health and healthcare equity. Supports targeted interventions for at-risk populations using AI-driven remote monitoring.
Real-World Integration Potential with Electronic Health Records (EHR)	AI models in nephrology were proposed but not systematically confirmed for clinical integration [8, 30-32].	Findings suggest XGBoost and Deep Learning models can be integrated into hospital EHR systems for: Automated UTI risk stratification. Personalized treatment recommendations (*e.g.*, antibiotic prophylaxis). Reducing unnecessary imaging and invasive interventions.	Bridges the gap between AI research and real-world clinical application. AI-based models can enhance pediatric nephrology decision support systems.

Several limitations must be acknowledged; all patient data were collected from a private pediatric surgery clinic, which may limit the generalizability of the findings. A multi-center study with diverse patient populations would enhance the external validity of the model. This study relies on retrospective data, which may introduce selection bias and limit the ability to set up causal relationships. A prospective validation with real-time data collection would strengthen the study's predictive utility. Although 211 patients were included, larger datasets are necessary to improve model robustness, especially for deep-learning approaches. A small sample size can lead to overfitting, particularly in complex models like neural networks. While the study identifies rural residence as a key predictor, other socioeconomic factors (*e.g.*, parental education and household income) were not fully analyzed. Future research should incorporate comprehensive socioeconomic variables to refine risk stratification further. The AI models were trained and assessed on a single dataset without external validation on an independent patient cohort. Validation on larger, geographically diverse datasets is needed to confirm model reliability. While SHAP and LIME were used to improve interpretability, deep learning models are still black-box systems, making clinical decision-making challenging. Further research is needed to enhance explainable AI (XAI) frameworks for real-world pediatric applications. The AI models have not yet been integrated into real-world clinical workflows or DHR systems. Pilot testing in clinical settings is crucial for evaluating usability, acceptance, and clinical impact. Data entry and coding errors in the electronic dataset could introduce bias. Automated data verification mechanisms should be incorporated in future studies. The study does not assess the impact of AI-driven predictions on antibiotic prescribing patterns. Future research should investigate whether AI-guided decision-making enhances antibiotic use and reduces antibiotic resistance.

CONCLUSION

This study sets up a robust, AI-driven framework for predicting recurrent UTIs in pediatric patients, proving clear advantages over traditional statistical methods. By utilizing machine learning and deep learning, this model enhances early risk identification, facilitates personalized treatment strategies, and promotes equitable healthcare access. Given its high predictive performance, clinical interpretability, and potential for DHR integration, this AI-based approach stands as a promising advancement in pediatric nephrology and urology, paving the way for improved outcomes and more efficient healthcare delivery.

AUTHORS' CONTRIBUTIONS

The authors confirm their contribution to the paper as follows: M.A.: Study conception and design; M.K.: Analysis and interpretation of results; M.R.: Validation; S.M.K.: Draft manuscript. All authors reviewed the results and approved the final version of the manuscript.

LIST OF ABBREVIATIONS

UTI	=	Urinary Tract Infection
VUR	=	Vesicoureteral Reflux
CKD	=	Chronic Kidney Disease
AI	=	Artificial Intelligence
ML	=	Machine Learning
DL	=	Deep Learning
CNN	=	Convolutional Neural Network
SVM	=	Support Vector Machine
NLP	=	Natural Language Processing
DHR	=	Digital Health Record
EHR	=	Electronic Health Record
VCUG	=	Voiding Cystourethrography
DMSA	=	Dimercaptosuccinic Acid (used in renal scan)
LR	=	Logistic Regression
RF	=	Random Forest
LIME	=	Local Interpretable Model-Agnostic Explanations
OR	=	Odds Ratio
CI	=	Confidence Interval
WBC	=	White Blood Cell (count)
CRP	=	C-Reactive Protein
PPV	=	Positive Predictive Value
NPV	=	Negative Predictive Value
AUC-ROC	=	Area Under the Receiver Operating Characteristic Curve
IQR	=	Interquartile Range
SD	=	Standard Deviation
TRIPOD	=	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
STROBE	=	Strengthening the Reporting of Observational Studies in Epidemiology
HIPAA	=	Health Insurance Portability and Accountability Act (U.S. data protection law)
GDPR	=	General Data Protection Regulation (EU data protection law)
MICE	=	Multiple Imputation by Chained Equations

ETHICAL STATEMENT

This study was a retrospective analysis of pre-existing, anonymized clinical data collected over 15 years. All patient information was fully de-identified before analysis, and no direct patient contact, intervention, or new data collection occurred during the course of the study. Following the international ethical guidelines (including the Declaration of Helsinki, 2013 revision) and national research policies, retrospective studies using anonymized data without potential risk to participants are exempt from formal ethical review.

CONSENT FOR PUBLICATION

STANDARDS OF REPORTING

STROBE guidelines were followed.

AVAILABILITY OF DATA AND MATERIALS

The data supporting the findings of the article will be available from the corresponding author [M.A] upon reasonable request.

FUNDING

CONFLICT OF INTEREST

The author(s) declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

REFERENCES

[1] Hoberman A, Greenfield SP, Mattoo TK, *et al.* Antimicrobial prophylaxis for children with vesicoureteral reflux. N Engl J Med 2014; 370(25): 2367-76.
http://dx.doi.org/10.1056/NEJMoa1401811 PMID: 24795142

[2] Montini G, Tullus K, Hewitt I. Febrile urinary tract infections in children. N Engl J Med 2011; 365(3): 239-50.
http://dx.doi.org/10.1056/NEJMra1007755 PMID: 21774712

[3] Salo J, Ikäheimo R, Tapiainen T, Uhari M. Childhood urinary tract infections as a cause of chronic kidney disease. Pediatrics 2011; 128(5): 840-7.
http://dx.doi.org/10.1542/peds.2010-3520 PMID: 21987701

[4] Shaikh N, Morone NE, Bost JE, Farrell MH. Prevalence of urinary tract infection in childhood: A meta-analysis. Pediatr Infect Dis J 2008; 27(4): 302-8.
http://dx.doi.org/10.1097/INF.0b013e31815e4122   PMID: 18316994

[5] Chowdhury Adiba Tabassum, Salam Abdus, Naznine Mansura, Abdalla Da'ad, Erdman Lauren, Chowdhury Muhammad E H, *et al.* Artificial intelligence tools of recent advances. Diagnostics 2024; 14(18): 2059.
http://dx.doi.org/10.3390/diagnostics14182059

[6] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. 31st Conference on Neural Information Processing Systems (NIPS 2017). Long Beach, CA, USA, 2017, pp.1-10

[7] Shaikh N, Craig JC, Rovers MM, *et al.* Identification of children and adolescents at risk for urinary tract infection recurrence: A systematic review and meta-analysis. JAMA Pediatr 2014; 170(9): 848-54.
http://dx.doi.org/10.1001/jamapediatrics.2016.1181   PMID: 27455161

[8] Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. Nat Med 2019; 25(1): 44-56.
http://dx.doi.org/10.1038/s41591-018-0300-7 PMID: 30617339

[9] Choi Eom Ji, Lee Min Ju, Park Sin-Ae, Lee Oh-Kyung. Predictors of high-grade vesicoureteral reflux in Children with febrile urinary tract infections. Child Kidney Dis 2017; 21(2): 136-41.
http://dx.doi.org/10.3339/jkspn.2017.21.2.136

[10] Storme O, Tirán Saucedo J, Garcia-Mora A, Dehesa-Dávila M, Naber KG. Risk factors and predisposing conditions for urinary tract infection. Ther Adv Urol 2019; 11: 1756287218814382.
http://dx.doi.org/10.1177/1756287218814382 PMID: 31105772

[11] Jiang J, Chen X-Y, Guo H. Clinical characteristics and nomogram model for predicting the risk of recurrent urinary tract infection in children. Sci Rep 2024; 14: 76901.
http://dx.doi.org/10.1038/s41598-024-76901-0 PMID: 39455869

[12] Wheeler D, Vimalachandra D, Hodson EM, Roy LP, Smith G, Craig JC. Antibiotics and surgery for vesicoureteric reflux: A meta-analysis of randomised controlled trials. Arch Dis Child 2003; 88(8): 688-94.
http://dx.doi.org/10.1136/adc.88.8.688 PMID: 12876164

[13] Keren R, Shaikh N, Pohl H, *et al.* Risk factors for recurrent urinary tract infection and renal scarring. Pediatrics 2015; 136(1): e13-21.
http://dx.doi.org/10.1542/peds.2015-0409 PMID: 26055855

[14] Wennerström M, Hansson S, Jodal U, Sixt R, Stokland E. Renal function 16 to 26 years after the first urinary tract infection in childhood. Arch Pediatr Adolesc Med 2000; 154(4): 339-45.
http://dx.doi.org/10.1001/archpedi.154.4.339 PMID: 10768669

[15] Foxman B. Urinary tract infection syndromes: Occurrence, recurrence, bacteriology, risk factors, and disease burden. Infect Dis Clin North Am 2014; 28(1): 1-13.
http://dx.doi.org/10.1016/j.idc.2013.09.003 PMID: 24484571

[16] Mattoo TK, Shaikh N, Nelson CP. Contemporary management of urinary tract infection in children. Pediatrics 2021; 147(2): e2020012138.
http://dx.doi.org/10.1542/peds.2020-012138

[17] Goodfellow I, Bengio Y, Courville A. Deep learning. MIT Press 2016.

[18] Kerth JL, Hagemeister M, Bischops AC, *et al.* Artificial intelligence in the care of children and adolescents with chronic diseases: A systematic review. Eur J Pediatr 2024; 184(1): 83.
http://dx.doi.org/10.1007/s00431-024-05846-3 PMID: 39672974

[19] van Smeden M, Reitsma JB, Riley RD, Collins GS, Moons KGM. Clinical prediction models: Diagnosis *versus* prognosis. J Clin Epidemiol 2021; 132: 142-5.
http://dx.doi.org/10.1016/j.jclinepi.2021.01.009 PMID: 33775387

[20] Vollmer S, Mateen BA, Bohner G, *et al.* Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. BMJ 2020; 368: l6927.
http://dx.doi.org/10.1136/bmj.l6927

[21] Tomašev N, Glorot X, Rae JW, *et al.* A clinically applicable approach to continuous prediction of future acute kidney injury. Nature 2019; 572(7767): 116-9.
http://dx.doi.org/10.1038/s41586-019-1390-1 PMID: 31367026

[22] Koyner JL, Carey KA, Edelson DP, Churpek MM. The development of a machine learning inpatient acute kidney injury prediction model. Crit Care Med 2018; 46(7): 1070-7.
http://dx.doi.org/10.1097/CCM.0000000000003123   PMID: 29596073

[23] Targeted workup after initial febrile urinary tract infection: Using a novel machine learning model to identify children most likely to benefit from VCUG. J Urol 2019; 202(1): 144-52.
http://dx.doi.org/10.1097/JU.0000000000000186 PMID: 30810465

[24] Urinary tract infection in under 16s: Diagnosis and management. London: National Institute for Health and Care Excellence (NICE) 2018.

[25] Grumbach K, Lucey CR, Johnston SC. Transforming from centers

of learning to learning health systems: The challenge for academic health centers. JAMA 2014; 311(11): 1109-10.
http://dx.doi.org/10.1001/jama.2014.705 PMID: 24643597

[26]   Hawkins DM, Basak SC, Mills D. Assessing model fit by cross-validation. J Chem Inf Comput Sci 2003; 43(2): 579-86.
http://dx.doi.org/10.1021/ci025626i PMID: 12653524

[27]   Hosmer DW, Lemeshow S, Sturdivant RX. Applied logistic regression. Wiley Series in Probability and Statistics 2013.
http://dx.doi.org/10.1002/9781118548387

[28]   Shen L, An J, Wang N, Wu J, Yao J, Gao Y. Artificial intelligence and machine learning applications in urinary tract infections identification and prediction: A systematic review and meta-analysis. World J Urol 2024; 42(1): 464.
http://dx.doi.org/10.1007/s00345-024-05145-4 PMID: 39088072

[29]   Rajkomar A, Oren E, Chen K, *et al.* Scalable and accurate deep learning with electronic health records. NPJ Digit Med 2018; 1(1): 18.

http://dx.doi.org/10.1038/s41746-018-0029-1 PMID: 31304302

[30]   Roberts Kenneth B, Downs Stephen M, Finnell S Maria E, *et al.* Reaffirmation of AAP clinical practice guideline: The diagnosis and management of the initial urinary tract infection in febrile infants and young children 2–24 months of age. Pediatrics 2016; 138(6): e20163026.
http://dx.doi.org/10.1542/peds.2016-3026

[31]   Chen Y, Ge XH, Yu Q, *et al.* Prediction model for urinary tract infection in pediatric urological surgery patients. Front Public Health 2022; 10: 888089.
http://dx.doi.org/10.3389/fpubh.2022.888089 PMID: 35812501

[32]   Ribeiro MT, Singh S, Guestrin C. "Why should i trust you?": Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA, 13 August 2016,pp. 1135-1144.
http://dx.doi.org/10.1145/2939672.2939778