Distribution Agreement

Signature:

______________________________ ______________

Jiaoan Chen Date

Neural Network and Calculator for Predicting Breakthrough Febrile Urinary Tract Infections in Children with Primary Vesicoureteral Reflux

By

Jiaoan Chen
Master of Science in Public Health

Biostatistics and Bioinformatics

Traci Leong
(Thesis Advisor)

Yi-an Ko
(Reader)

Neural Network and Calculator for Predicting Breakthrough Febrile Urinary Tract Infections in Children with Primary Vesicoureteral Reflux

By

Jiaoan Chen

B.E.
Shandong University of Science and Technology
2018

Thesis Committee Chair: Traci Leong, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Biostatistics and Bioinformatics
2020

Abstract

Neural Network and Calculator for Predicting Breakthrough Febrile Urinary Tract Infections in Children with Primary Vesicoureteral Reflux
By Jiaoan Chen

Background: Factors that influence the decision to surgically correct VUR include risk of developing new renal parenchymal scarring, further risk of breakthrough fUTI, and the likelihood of spontaneous VUR resolution. Improved identification of children at risk for breakthrough fUTI could help make decisions regarding whether or not to surgically correct VUR. We constructed three models to predict breakthrough fUTI and evaluated the accuracy of these models.

Methods and Materials: Medical records of 384 children diagnosed with primary VUR in whom detailed voiding cystourethrogram (VCUG) and clinical data were documented between 1984 and 2010 were reviewed. The variables involved in our analyses included age, gender, percentage of PBC at VUR onset, VUR grade (0-2, 3, 4-5), laterality, history of BBD, number of UTIs prior to VUR diagnosis (≥ 2 vs. <2), history of fUTI, VUR onset (filling or voiding) and dilating VUR (0-2, 3-5). The data was randomized into a training set of 288 for model creation and a validation (testing) set of 96, following the 75/25-splitting rule. The data was modeled with logistic regression, neural network, and random forest. Receiver operating characteristic (ROC) area was utilized to assess the performance of the model.

Results: A total of 384 patients with primary reflux were recruited in the study, with 64 male patients (16.67%) and 320 female patients (83.33%). The number of patients developing the outcome of breakthrough fUTI was 128 (33.33%). Gender, percentage of PBC at VUR onset, VUR grade, bilateral VUR, BBD, history of fUTI were significantly associated with breakthrough fUTI. The model that best fit the data and had the highest discrimination ability was a one-hidden node neural network model, with an AUC of 0.756.

Conclusion: Our neural network model, using multiple variables, predicts breakthrough fUTI on an individual basis with an AUC of 76%. Our prognostic calculator based on the neural network model will provide a useful tool for users to conveniently get a prediction and a probability of developing breakthrough fUTI. Such prognostic information can assist in clinical decision-making and provide useful information in reflux management.

Neural Network and Calculator for Predicting Breakthrough Febrile Urinary Tract Infections in Children with Primary Vesicoureteral Reflux

By

Jiaoan Chen

B.E.
Shandong University of Science and Technology
2018

Thesis Committee Chair: Traci Leong, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Biostatistics and Bioinformatics
2020

CONTENTS

1. INTRODUCTION

Vesicoureteral reflux (VUR) is a condition in which urine flows backward from the bladder to one or both ureters and sometimes to the kidney. This condition is most common in infants and young children, with an estimated overall prevalence of 7% among children presenting with fever (Shaikh et al., 2008). Some children resolve VUR spontaneously by themselves as they get older, but some children cannot resolve the reflux. Among these children without resolution, recurrent breakthrough febrile urinary tract infection (fUTI) and renal scarring can occur, with possible subsequent sequelae such as proteinuria, hypertension and chronic kidney disease (Peters and Rushton, 2010). To avoid these complications, a child with primary VUR needs prophylactic administration of antibiotic, endoscopic injection, or even an antireflux surgery to manually correct it (Elder et al., 1997). Factors that influence the decision to surgically correct VUR include risk of developing new renal parenchymal scarring, further risk of breakthrough fUTI, and the likelihood of spontaneous VUR resolution (Arlen et al., 2016). Thus, it is very important to have an estimate of the risk of developing breakthrough fUTI to allow for optimal management of VUR for children. This can include, to wait until spontaneous resolution of VUR or to surgically correct VUR in preventing future illness or injury.

Several statistical methods have been used to predict breakthrough fUTI, among which logistic regression is the most commonly used one. Previous studies have shown that the risk factors related to breakthrough fUTI include high-grade VUR, gender, history of fUTI, etc., and a multivariable logistic regression model could successfully predict breakthrough fUTI with reasonable accuracy (Hidas et al., 2015). With the increasing popularity of machine learning methods, some research groups tried using machine learning methods to conduct a prediction for breakthrough fUTI aiming at having a higher prediction accuracy (Wu et al., 2019). Classification and regression trees (CART) are machine-learning methods for classification or constructing prediction models from data. Random forest is just an advanced version of CART, which operates by constructing numerous decision trees at training and outputting the class which is the mode of classes or the mean prediction of individual trees (Krzywinski and Altman, 2017). The random forest model might work well since a multitude of relatively uncorrelated models (trees) operating as a whole would outperform any of the individual models, generating a more accurate prediction than that of any individual tree (Ho, 1998). However, the random forest model did not perform better than the logistic regression model in predicting breakthrough fUTI, with relatively low prediction accuracy (Alexander et al., 2015). Thus, additional machine learning methods are being explored, among which the neural network is attractive for its high prediction accuracy and good performance in healthcare.

Artificial neural networks (ANNs) are computing systems based on mathematical models that, unlike traditional computing models, have a structure and operation that resemble that of the human brain. ANNs have been used in healthcare for several decades, with a purpose to transform huge amounts of raw data into useful decisions for treatment and care (Begg et al., 2006). Artificial neural networks could be used to interpret clinical data and perform accurate diagnosis of disease or prediction of future risk, and compared to other statistical models, they were shown to be more accurate and efficient in diagnosis and prediction (Abdul Basit Shaikh et al., 2014). Also, it has been demonstrated that ANNs can accurately predict the chance of spontaneous VUR resolution, and thus, could avoid the unnecessary exposure of children with primary reflux in x-ray radiation examination, or allow urologists to aid in the decision-making process of VUR treatment (Seckiner et al., 2008). While other studies have successfully demonstrated the application of using ANNs in healthcare and spontaneous VUR resolution, to date, few studies have explored the use of ANN for breakthrough fUTI prediction. As improved identification of children at risk for breakthrough fUTI could help make decisions regarding whether or not to surgically correct VUR, this study addresses the need for more accurate prediction of breakthrough fUTI using neural network modeling.

In this study, we demonstrated the performance of a neural network model in predicting breakthrough fUTI in comparison to logistic regression and random forest models. We also used the neural network model to create a prognostic calculator to allow input variables to be entered to calculate the probability of a child developing breakthrough fUTI. Our calculator would contain two parts including a user interface (UI) controlling the outlook of the web page and a server controlling the logic, which makes it easier and more convenient for other users to get the prediction results by just simply entering a patient's characteristics. Together, these findings aim to use the neural network modeling to obtain a more accurate prediction of breakthrough fUTI, and thus, to help urologists make decisions and children with primary reflux get rid of infection and disease.

2. METHODS AND MATERIALS

2.1 Dataset

Approval of using the data was obtained from the University of Iowa Hospitals & Clinics IRB (2014-04766). Medical records of 384 children diagnosed with primary VUR in whom detailed voiding cystourethrogram (VCUG) and clinical data were documented between 1984 and 2010 were reviewed. The recorded findings from cystogram included reflux grade, laterality, and bladder volume at VUR onset. VUR grade was determined using the International Reflux Study classification system (Lebowitz et al., 1985). Bladder volume at VUR onset was normalized by age predicted bladder capacity (PBC) for all patients using the equation, PBC = (age + 2) × 30 mL (age in years) (Arlen et al., 2016). All children with primary reflux were given prophylactic antibiotics and followed with cystograms annually until spontaneous resolution or operative repair. Children without continuous antibiotic prophylaxis or with secondary VUR were excluded from this study.

Clinical variables including age (at diagnosis), gender, number of urinary tract infections (UTIs) prior to VUR diagnosis, history of fUTI, and the presence of bladder and bowel dysfunction (BBD) were recorded. Febrile urinary tract infection (fUTI) was defined as a positive urine culture of >100,000 colony-forming units (CFU) of a single organism associated with a body temperature of >101.4 F (38.5 C) documented in the medical record. BBD was defined as incontinence episodes more than expected for age or requiring a prescription of anticholinergic medication or laxatives (Arlen et al., 2016).

2.2 Data Preparation

The variables involved in our analyses included age, gender, percentage of PBC at VUR onset, VUR grade (0-2, 3, 4-5), laterality, history of BBD, number of UTIs prior to VUR diagnosis (≥2 vs. <2), history of fUTI, VUR onset (filling or voiding) and dilating VUR (0-2, 3-5). Variables other than age, percentage of PBC at VUR onset and VUR grade were converted into a binary numerical code. Data at the ureter level rendered both left and right VUR grade and VUR onset. To keep all data at the patient level, maximum grade and worst level of onset were only considered. VUR grade was initially recorded as 6 groups: 0, 1, 2, 3, 4, 5, but we combined 0-2 to a single group (low grade) and 4-5 to another group, resulting three groups which were 0-2, 3, 4-5. Dilating VUR was defined as a VUR grade of 3-5. This variable was different from VUR grade and was created to see whether high grade VUR had influence on breakthrough fUTI. The binary outcome of interest was whether a patient developed breakthrough fUTI or not. The data was then randomized into a training set of 288 for model creation and a validation (testing) set of 96, following the 75/25-splitting rule.

2.3 Statistical Methods

Statistical analyses were conducted using R version 3.6.0. Descriptive statistics for all the variables were reported, with the mean and standard deviation for continuous variables and the frequency and percentage for categorical variables. Univariate logistic regression analysis for breakthrough fUTI was first performed for all patients and then among patients in the training dataset. Odds ratio, 95% confidence interval, and p-value reflected the results of hypothesis tests on the relationship between any variable and breakthrough fUTI. The significant level was set to be 0.05.

The data was modeled with logistic regression, neural network, and random forest after investigating several model types using the training dataset. Input features included age, gender, percentage of PBC at VUR onset, VUR grade, laterality, history of BBD, number of UTIs prior to VUR diagnosis, history of fUTI and VUR onset. We used VUR grade instead of dilating VUR in modeling because grade 3 was significantly associated with breakthrough fUTI and the effects of grade 3 ($p = 0.0494$) and grade 4-5 ($p = 0.510$) for breakthrough fUTI were different. Multivariable logistic regression, random forest and neural network models were computed, and each model's accuracy was evaluated using the testing dataset. Receiver operating characteristic (ROC) area was utilized to assess the performance of the model, where 1 (on a scale of 0-1) indicated a perfect measure of discrimination. It was calculated using the statistical method described by Wickens (Wickens 2002). An accuracy table was also generated to compare the performance of these three models, which included sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV).

For the neural network model, the package we used was "neuralnet". Min-Max normalization was done to scale continuous variables down since neural networks used activation functions between -1 and +1, and to ensure that a continuous variable didn't dominate. Categorical variables were converted into dummy variables to be useful for the package. Different parameters including hidden node, stepmax and threshold were tuned to determine the optimal values used to optimize our model.

2.4 Shiny Calculator

A shiny calculator was created using the constructed neural network model because of relatively high prediction accuracy, for allowing other users to get prediction results of breakthrough fUTI by simply entering a child's characteristics. Users are supposed to choose a value for gender, bilaterality, BBD, number of UTIs, history of fUTI, VUR grade and VUR onset, and type a numeric value for age and percentage of PBC at VUR onset. A prediction value (0 for no & 1 for yes) as well as a probability will be generated after clicking the submit button. A figure was generated to show the interface of our prognostic calculator.

3. RESULTS

3.1 Descriptive Analysis

Descriptive statistics for all the variables were provided in **Table 1**. The data used in this study contained 10 explanatory variables that were considered in the prediction of breakthrough fUTI. A total of 384 patients with primary reflux were recruited in the study, with 64 male patients (16.67%) and 320 female patients (83.33%). The mean age of our patients was 3.90 $\pm$ 3.05 years, and the mean percentage of predicted bladder capacity at vesicoureteral reflux onset was 51.07 $\pm$ 38.84 %. VUR was grade 0-2 (n = 248, 64.58%), grade 3 (n = 107, 27.86%) and grade 4-5 (n = 29, 7.55%). The number of patients having dilating VUR was 136 (35.42%). One hundred and seventy-eight patients (46.35%) had bilateral VUR. Ninety-eight children (25.52%) had documented BBD. Two hundred and seventy-eight patients (72.40%) had <2 prior UTI, and 106 patients (27.60%) had ⩾2 UTIs prior to VUR diagnosis. Two hundred and thirty-nine patients (62.24%) had a history of fUTI. Three hundred and thirty-one patients (86.20%) had filling VUR onset, and 53 patients had voiding VUR onset (13.80%). The number of patients developing the outcome of breakthrough fUTI was 128 (33.33%). Median follow-up time was 24 months (interquartile range 12 to 52 months).

3.2 Univariate Analysis

Univariate logistic regression analyses for breakthrough fUTI were performed. Bivariate associations between each variable and developing breakthrough fUTI were summarized in **Table 2a** and **Table 2b**, using the whole and training datasets, respectively. The results of these two datasets were similar to each other as expected. Remaining results will be based on either training or testing data, not entire dataset. Breakthrough fUTI was significantly more common in females than males (OR = 3.452, 95% CI = 1.573-8.706, p = 0.00398). There was a significant association between high percentage of PBC at VUR onset with decreased risk of breakthrough fUTI (OR = 0.984, 95% CI = 0.975-0.993, p = 0.000556). VUR grade 3 was significantly associated with breakthrough fUTI (OR = 1.727, 95% CI = 0.999-2.979, p = 0.0494). Bilateral VUR (OR = 1.961, 95% CI = 1.197-3.235, p = 0.00786), BBD (OR = 2.325, 95% CI = 1.345-4.025, p = 0.00248), and history of fUTI (OR = 2.073, 95% CI = 1.233-3.556, p =

0.0068) were also significantly associated with an increased risk of breakthrough fUTI.

3.3 Multivariate Logistic Regression Model and Random Forest Model

A multivariate logistic regression model was constructed to predict the likelihood of developing breakthrough fUTI. Summary statistics for the multivariate logistic regression model were presented in **Table 3**. Percentage of PBC at VUR onset ($p = 0.019$) and BBD ($p = 0.00938$) were the variables contributing significantly in predicting breakthrough fUTI. We tested the risk model in a validation (testing) dataset and generated a ROC curve to assess the performance of the model with the area under ROC curve (AUC). Mean ± SD predicted probability of developing breakthrough fUTI was 32.2% ± 15.3% (range 7% to 70%). As shown in **Figure 1**, ROC curve with its 95% confidence interval demonstrated a good discrimination between presence and absence of breakthrough fUTI of the logistic regression model (AUC 0.74).

A random forest model was also constructed with breakthrough fUTI as the outcome. After tuning several parameters, we built the final random forest model with 500 trees and 3 variables tried at each split. We also tested the risk model in a validation (testing) dataset and generated a ROC curve to assess the performance of the model with AUC. Mean ± SD predicted probability of developing breakthrough fUTI was 50.0% ± 25.3% (range 4% to 96%). As shown in **Figure 2**, ROC curve with its 95% confidence interval demonstrated fair discrimination between presence and absence of breakthrough fUTI of the random forest model (AUC 0.67), which was less than that of the logistic regression model.

3.4 Neural Network Model

A neural network model with 1 hidden node was constructed after several attempts of the number of hidden nodes. **Figure 3** shows the structure of our neural network model in details. We also used a validation (testing) dataset to test our risk model and generated a ROC curve to assess the performance of the model with AUC. Mean ± SD predicted probability of developing breakthrough fUTI was 33.3% ± 15.3% (range 1% to 61%). As shown in **Figure 4**, ROC curve with its 95% confidence interval demonstrated a good discrimination between presence and absence of breakthrough fUTI of the neural network model (AUC 0.76), which was slightly higher than that of the previous two models.

3.5 Comparison of Performance for Different Models

For all the three models mentioned above, the model that best fit the data and had the highest discrimination ability was the neural network model, with the largest AUC which was 0.756 (**Table 4**). The model that had the second highest discrimination ability was the multivariate logistic regression model, which was better than the random forest model (AUC 0.736 vs 0.669, p = 0.417) and was not significantly worse than the neural network model (p = 0.798). **Figure 5** shows how well the predictive variables fit these three models, and the ROC curves demonstrated that the neural network model had better performance than the other two models. Setting the threshold specifically to be 0.5, some measures of prediction accuracy including sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) could be calculated for different models (**Table 5**). The higher the values of these four measures were, the better the model would be. We could see from the table that the multivariate logistic regression model and the neural network model were much better than the random forest model in predicting breakthrough fUTI. The values for sensitivity, specificity, PPV and NPV were all relatively high for the neural network model, indicating relatively good prediction accuracy.

Overall, the neural network model was shown to have the best performance among all of the three models in predicting breakthrough fUTI, which contributes to having a more accurate prediction with a patient's risk factors.

3.6 Shiny Calculator

We developed the prognostic shiny calculator, a web-based application to allow other users making a prediction of breakthrough fUTI from a patient's characteristics. **Figure 6** shows the interface of our calculator that users would be able to access. We termed this interface a prognostic calculator since the 9 variables could be input and an individualized prediction would be generated. As shown in the figure, users will get a prediction of breakthrough fUTI, where 1 indicates developing the infection and 0 indicates not developing, as well as a probability of developing it. Based on the prediction results from this calculator, people can make decisions regarding whether or not to surgically repair VUR for children with primary reflux.

4. CONCLUSION

According to our study, compared with the multivariate logistic regression and random forest models, the neural network model has a better performance in predicting breakthrough fUTI. Our neural network model, using multiple variables, predicts breakthrough fUTI on an individual basis with an AUC of 76%. Our prognostic calculator based on the neural network model will provide a useful tool for users to conveniently get a prediction and a probability of developing breakthrough fUTI. Such prognostic information can assist in clinical decision-making and provide useful information in reflux management.

5. DISCUSSION

Vesicoureteral reflux (VUR) is a urologic condition that is most common in infants and young children, with an estimated overall prevalence of 7% among children presenting with fever (Shaikh et al., 2008). Further risk of breakthrough fUTI is a factor that influence the decision to surgically correct VUR (Arlen et al., 2016). Since increased risk of breakthrough fUTI can make children with primary VUR suffer potential danger to develop diseases, it is important to identify a subpopulation of children who are at risk for developing breakthrough fUTI so that optimal management of VUR could be done. We tried three methods to conduct the prediction of breakthrough fUTI. For all the models constructed to predict breakthrough fUTI, we found that a neural network model has the best performance and can give the most accurate prediction.

Previous studies showed that there were several risk factors that were associated with developing breakthrough fUTI, such as VUR grade, history of fUTI, gender and BBD (Hidas et al., 2015). They built some statistical models using those identified risk factors to get a prediction of breakthrough fUTI, but mostly the prediction accuracy was not very high (under 70%). Though neural network was applied widely and successfully in healthcare, few of them tried constructing a neural network model to predict the breakthrough fUTI.

Through univariate analyses, we found that factors that were associated with breakthrough fUTI included gender, percentage of PBC at VUR onset, VUR grade, bilateral VUR, BBD and history of fUTI, which was consistent with previous studies. However, there were still some other factors that were shown to have possible correlation with breakthrough fUTI, so we incorporated all known risk factors to construct our prediction model. We demonstrated that a multivariate logistic regression model can provide a reasonably accurate prediction, but a random forest model performed not very well. The comparison of all the models (comparison of AUCs, sensitivities, etc.) demonstrated that a neural network model can provide a prediction with higher accuracy, suggesting that there are some possibilities to increase the prediction accuracy when predicting breakthrough fUTI.

Our data, however, does not include the patient's follow up time which might be associated with breakthrough fUTI. It is possible that longer follow up time could be related to the breakthrough fUTI incidence. Also, though we can get a relatively accurate prediction through constructing our model, the accuracy does not achieve our expectations. We might try some other methods in the future to see whether we can further improve our prediction accuracy. Finally, our patient cohort is from a single institution, but acquiring patient data from multiple institutions would help validate our constructed model and provide more reliable results.

Overall, despite the limitations, our prognostic computational model predicts breakthrough fUTI on an individual basis with relatively high accuracy (AUC 0.76). Our prognostic calculator provides a useful platform for users to conveniently get a prediction and a probability of developing breakthrough fUTI. These will help urologists make decisions regarding whether to surgically repair VUR, and thus give children with primary reflux the best management and help them get rid of infection and disease.

6. REFERENCE

Shaikh N, Morone N. E, Bost J. E and Farrell M. H. 2008. “Prevalence of urinary tract infection in childhood: a meta-analysis.” *Pediatr Infect Dis J*. 27: 302–308.

Peters C and Rushton H. G. 2010. “Vesicoureteral reflux associated renal damage: congenital reflux nephropathy and acquired renal scarring.” *J Urol.* 184: 265e73.

Elder J. S, Peters C. A, Arant B. S Jr, Ewalt D. H, Hawtrey C. E, Hurwitz R. S, et al. 1997. “Pediatric Vesicoureteral Reflux Guidelines Panel summary report on the management of primary vesicoureteral reflux in children.” *J Urol*. 157: 1846-51.

Arlen Angela M, Alexander Siobhan E, et al. 2016. “Computer model predicting breakthrough febrile urinary tract infection in children with primary vesicoureteral reflux.” *Journal of Pediatric Urology.* 12: 288.e1-288.e5.

Hidas Guy, Billimek John, et al. 2015. “Predicting the Risk of Breakthrough Urinary Tract Infections: Primary Vesicoureteral Reflux.” *The Journal of urology*. 194 (5): 1396-1401.

Charlotte Q. Wu, Andrew J. Kirsch, Elizabeth Traore, et al. (2019). “Vesicoureteral Reflux Index (VURx) is Independently Predictive of Breakthrough Febrile Urinary Tract Infection in Children with Primary Reflux.” Pediatric urology fall congress, Scottsdale, September 26-29, 2019.

Martin Krzywinski and Naomi Altman. 2017. “Classification and regression trees.” *Nature Methods*. 14: 757-758.

Tin Kam Ho. 1998. “The Random Subspace Method for Constructing Decision Forests.” *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 20 (8): 832-844.

Siobhan E. Alexander, Angela M. Arlen, Douglas W. Storm, Kathleen Kieran and Christopher S. Cooper. 2015. “Bladder Volume at Onset of Vesicoureteral Reflux is an Independent Risk Factor for Breakthrough Febrile Urinary Tract Infection.” *The Journal of Urology*. 193 (4): 1342-6.

Begg R, Kamruzzaman J and Sarker R. 2006. “Neural Networks in Healthcare: Potential and Challenges.” Hershey, United States: Idea Group Publishing.

Abdul Basit Shaikh, Muhammad Sarim, et al. 2014. “Artificial Neural Network: A Tool for Diagnosing Osteoporosis.” *Reasearch Journal of Recent Sciences.* 3 (2): 87-91.

Seckiner I, Seckiner S. U, et al. 2008. “The Use of Artificial Neural Networks in Decision Support in Vesicoureteral Reflux Treatment.” *Urol Int.* 80: 283-286.

Lebowitz R. L, Olbing H, et al. 1985. “International system of radiographic grading of vesicoureteric reflux.” *International Reflux Study in Children*. 15:105-9.

Wickens TD. 2002. “Elementary Signal Detection Theory.” New York: Oxford University Press.

7. TABLES AND FIGURES

Table 1. Descriptive statistics for all variables of patients with primary reflux (N = 384)

Characteristics	Mean (SD) / n (%)
Age	3.90 (3.05)
Gender	
Male	64 (16.67%)
Female	320 (83.33%)
Percentage of PBC at VUR onset	51.07 (38.84)
VUR Grade	
0-2	248 (64.58%)
3	107 (27.86%)
4-5	29 (7.55%)
Dilating VUR	
No	248 (64.58%)
Yes	136 (35.42%)
Bilateral VUR	
No	206 (53.65%)
Yes	178 (46.35%)
BBD	
No	286 (74.48%)
Yes	98 (25.52%)
History of UTI number	
< 2	278 (72.40%)
≥ 2	106 (27.60%)
History of fUTIs	
No	145 (37.76%)
Yes	239 (62.24%)
VUR Onset	
Filling	331 (86.20%)
Voiding	53 (13.80%)

VUR = Vesicoureteral Reflux, **PBC** = Predicted Bladder Capacity, **BBD** = Bladder and Bowel Dysfunction, **UTI** = Urinary Tract Infection, **fUTI** = Febrile Urinary Tract Infection

Table 2a. Univariate analysis results for all variables of all patients (N = 384)

Variable	N patients with BUTI	N patients without BUTI	Odds ratio (95% CI)	P-value
Age	128	256	1.047 (0.977, 1.122)	0.19
Gender				
Female	119	201	3.618 (1.807, 8.078)	0.000663*
Male	9	55	ref.	
Percentage of PBC at VUR onset	128	256	0.989 (0.981, 0.995)	0.00112*
VUR Grade				
0-2	74	174	ref.	
3	47	60	1.842 (1.151, 2.945)	0.0107*
4-5	7	22	0.748 (0.285, 1.748)	0.524
Dilating VUR				
Yes	54	82	1.548 (0.998, 2.400)	0.0505
No	74	174	ref.	
Bilateral VUR				
Yes	73	105	1.909 (1.245, 2.942)	0.00318*
No	55	151	ref.	
BBD				
Yes	47	51	2.332 (1.454, 3.747)	0.000443*
No	81	205	ref.	
History of UTI number				
≥2	47	59	1.937 (1.219, 3.077)	0.00505*
<2	81	197	ref.	
History of fUTI				
Yes	98	141	2.664 (1.669, 4.346)	5.74e-05*
No	30	115	ref.	
VUR Onset				
Filling	114	217	ref.	
Voiding	14	39	0.683 (0.346, 1.283)	0.252

* indicates significant (significant level = 0.05)
Dilating VUR defined as a grade of 3-5

Table 2b. Univariate analysis results for all variables of patients in the training dataset (N = 288)

Variable	N patients with BUTI	N patients without BUTI	Odds ratio (95% CI)	P-value
Age	96	192	1.038 (0.958, 1.124)	0.361
Gender				
Female	89	151	3.452 (1.573, 8.706)	0.00398*
Male	7	41	ref.	
Percentage of PBC at VUR onset	96	192	0.984 (0.975, 0.993)	0.000556*
VUR Grade				
0-2	56	128	ref.	
3	34	45	1.727 (0.999, 2.979)	0.0494*
4-5	6	19	0.722 (0.252, 1.811)	0.510
Dilating VUR				
Yes	40	64	1.429 (0.861, 2.366)	0.166
No	56	128	ref.	
Bilateral VUR				
Yes	55	78	1.961 (1.197, 3.235)	0.00786*
No	41	114	ref.	
BBD				
Yes	35	38	2.325 (1.345, 4.025)	0.00248*
No	61	154	ref.	
History of UTI number				
≥2	34	48	1.645 (0.965, 2.795)	0.066
<2	62	144	ref.	
History of fUTI				
Yes	69	106	2.073 (1.233, 3.556)	0.0068*
No	27	86	ref.	
VUR Onset				
Filling	86	163	ref.	
Voiding	10	29	0.654 (0.291, 1.363)	0.276

* indicates significant (significant level = 0.05)
Dilating VUR defined as a grade of 3-5

Table 3. A multivariate logistic regression model predicting breakthrough fUTI for patients in the training dataset

Variable	Coefficient estimate (SE)	Odds ratio (95% CI)	P-value
Age	-0.0247 (0.0486)	0.976 (0.885, 1.071)	0.611
Gender (Female) (Male as reference)	0.780 (0.498)	2.181 (0.859, 6.175)	0.117
Percentage of PBC at VUR onset	-0.0118 (0.00502)	0.988 (0.978, 0.997)	0.0190*
VUR Grade			
3	0.297 (0.309)	1.346 (0.731, 2.464)	0.337
4-5	-0.201 (0.580)	0.818 (0.248, 2.490)	0.729
(0-2 as reference)			
Bilateral VUR	0.510 (0.276)	1.666 (0.970, 2.874)	0.0650
BBD	0.795 (0.306)	2.214 (1.216, 4.050)	0.00938*
History of UTI number (≥2) (<2 as reference)	0.176 (0.303)	1.192 (0.655, 2.154)	0.561
History of fUTI	0.481 (0.306)	1.618 (0.893, 2.975)	0.116
VUR Onset (Voiding) (Filling as reference)	-0.191 (0.435)	0.826 (0.341, 1.906)	0.662

* indicates significant (significant level = 0.05)

Table 4. Test set ROC areas

Model	Test Set ROC Area
Multivariate Logistic Regression	0.736
Neural Network	0.756
Random Forest	0.669

Table 5. Accuracy table

	Multivariate Logistic Regression	Neural Network	Random Forest
TP (n)	9	10	7
FP (n)	5	6	7
TN (n)	59	58	57
FN (n)	23	22	25
Sensitivity (%)	28.13	31.25	21.88
Specificity (%)	92.19	90.63	89.06

PPV (%)	64.29	62.50	50.00
NPV (%)	71.95	72.50	69.51

TP = True Positive, **FP** = False Positive, **TN** = True negative, **FN** = False Negative, **PPV** = Positive predictive value, **NPV** = Negative Predictive Value

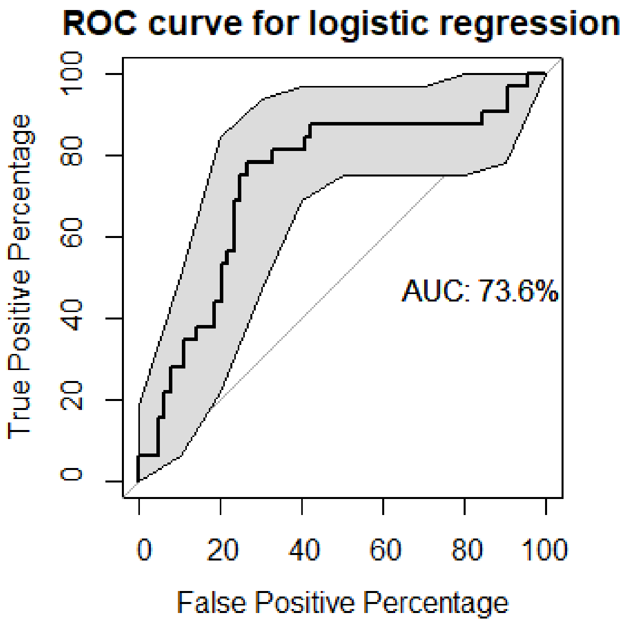

Figure 1. ROC curve for multivariate logistic regression model predicting breakthrough fUTI

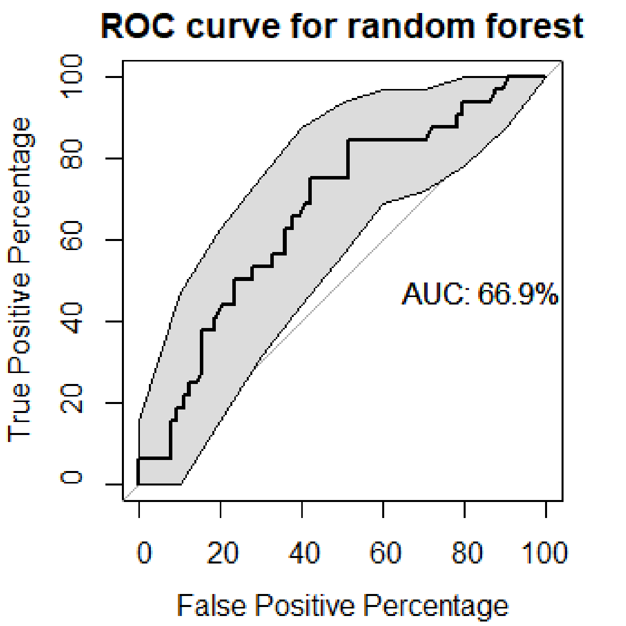

Figure 2. ROC curve for random forest model predicting breakthrough fUTI

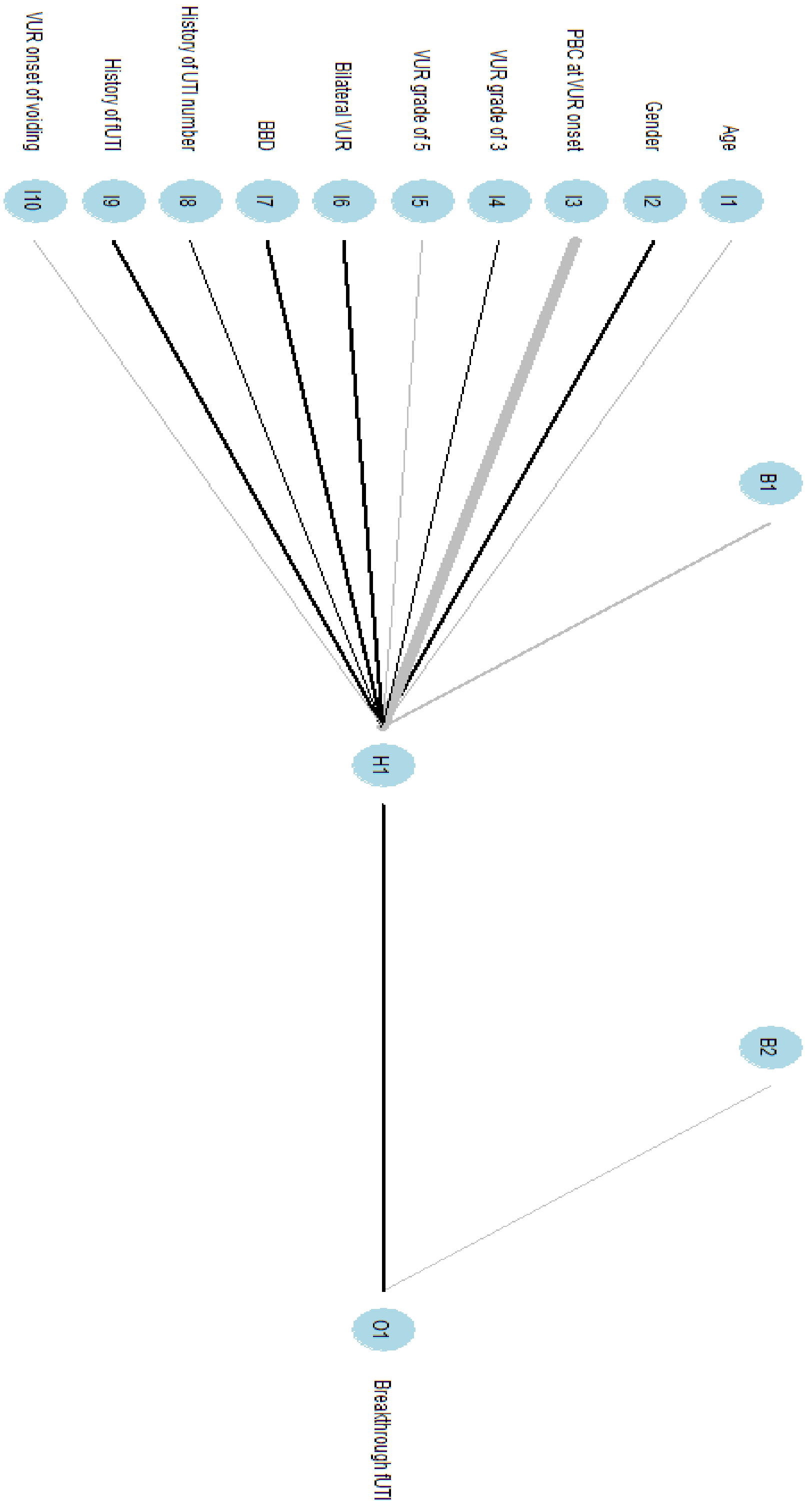

Figure 3. Structure of neural network model predicting breakthrough fUTI where B1 and B2 show the bias terms

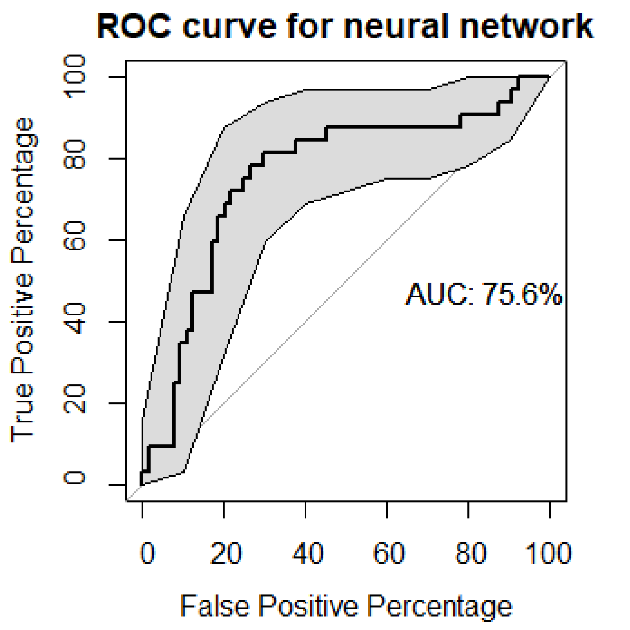

Figure 4. ROC curve for neural network model predicting breakthrough fUTI

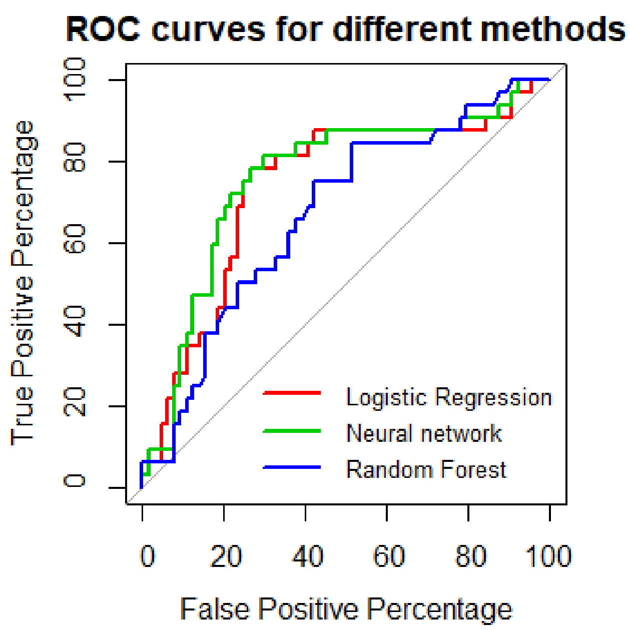

Figure 5. Comparison of ROC curves among LR, NN, and RF

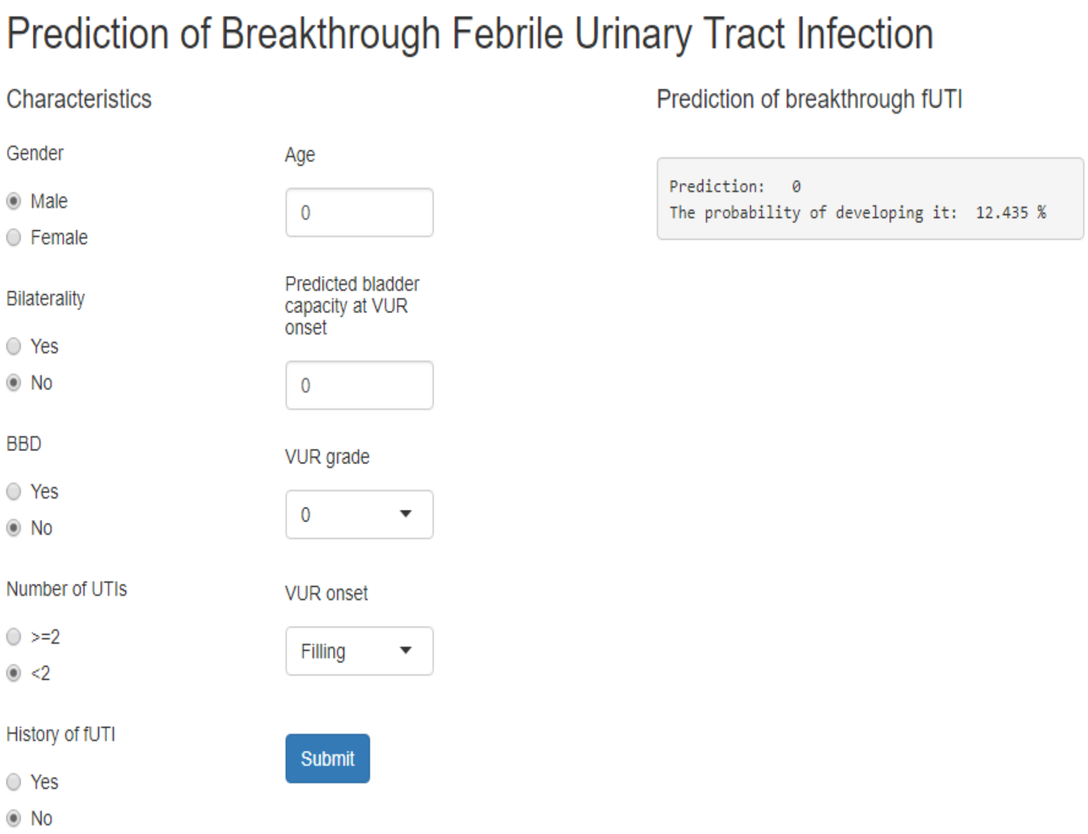

Figure 6. Interface for prognostic calculator predicting breakthrough fUTI