

Original Paper

Prediction of Antibiotic Resistance in Patients With a Urinary Tract Infection: Algorithm Development and Validation

Nevruz İlhanlı^{1,2*}, MSc; Se Yoon Park^{1,3,4*}, MD, PhD; Jaewoong Kim^{1,3}, MSc; Jee An Ryu¹, BA; Ahmet Yardımcı², PhD; Dukyong Yoon^{1,4,5}, MD, PhD

¹Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Yongin, Republic of Korea

²Department of Biostatistics and Medical Informatics, Akdeniz University, Antalya, Turkey

³Department of Hospital Medicine, Yongin Severance Hospital, Yonsei University College of Medicine, Yongin, Republic of Korea

⁴Center for Digital Health, Yongin Severance Hospital, Yonsei University Health System, Yongin, Republic of Korea

⁵Institute for Innovation in Digital Healthcare, Severance Hospital, Seoul, Republic of Korea

*these authors contributed equally

Corresponding Author:

Dukyong Yoon, MD, PhD

Department of Biomedical Systems Informatics

Yonsei University College of Medicine

363, Dongbaekjukjeon-daero

Yongin, 16995

Republic of Korea

Phone: 82 3151898450

Email: dukyong.yoon@yonsei.ac.kr

Abstract

Background: The early prediction of antibiotic resistance in patients with a urinary tract infection (UTI) is important to guide appropriate antibiotic therapy selection.

Objective: In this study, we aimed to predict antibiotic resistance in patients with a UTI. Additionally, we aimed to interpret the machine learning models we developed.

Methods: The electronic medical records of patients who were admitted to Yongin Severance Hospital, South Korea were used. A total of 71 features extracted from patients' admission, diagnosis, prescription, and microbiology records were used for classification. UTI pathogens were classified as either sensitive or resistant to cephalosporin, piperacillin-tazobactam (TZP), carbapenem, trimethoprim-sulfamethoxazole (TMP-SMX), and fluoroquinolone. To analyze how each variable contributed to the machine learning model's predictions of antibiotic resistance, we used the Shapley Additive Explanations method. Finally, a prototype machine learning-based clinical decision support system was proposed to provide clinicians the resistance probabilities for each antibiotic.

Results: The data set included 3535, 737, 708, 1582, and 1365 samples for cephalosporin, TZP, TMP-SMX, fluoroquinolone, and carbapenem resistance prediction models, respectively. The area under the receiver operating characteristic curve values of the random forest models were 0.777 (95% CI 0.775-0.779), 0.864 (95% CI 0.862-0.867), 0.877 (95% CI 0.874-0.880), 0.881 (95% CI 0.879-0.882), and 0.884 (95% CI 0.884-0.885) in the training set and 0.638 (95% CI 0.635-0.642), 0.630 (95% CI 0.626-0.634), 0.665 (95% CI 0.659-0.671), 0.670 (95% CI 0.666-0.673), and 0.721 (95% CI 0.718-0.724) in the test set for predicting resistance to cephalosporin, TZP, carbapenem, TMP-SMX, and fluoroquinolone, respectively. The number of previous visits, first culture after admission, chronic lower respiratory diseases, administration of drugs before infection, and exposure time to these drugs were found to be important variables for predicting antibiotic resistance.

Conclusions: The study results demonstrated the potential of machine learning to predict antibiotic resistance in patients with a UTI. Machine learning can assist clinicians in making decisions regarding the selection of appropriate antibiotic therapy in patients with a UTI.

(*JMIR Med Inform* 2024;12:e51326) doi: [10.2196/51326](https://doi.org/10.2196/51326)

KEYWORDS

antibiotic resistance; machine learning; urinary tract infections; UTI; decision support

Introduction

Urinary tract infection (UTI) refers to an infection that occurs in any part of the urinary system, including the kidneys, ureters, urinary bladder, urethra, and other auxiliary structures [1,2]. Globally, UTIs are the most prevalent type of infectious disease, with around 150-250 million cases occurring each year [3]. Considerable morbidity and mortality result from these infections [4]. Typically, the most effective treatment for UTIs is the administration of antibiotics [3]. However, inappropriate use of antibiotics can permanently affect the normal microbiota of the urinary tract system and lead to antibiotic resistance [5].

The antibiotic susceptibility test is commonly used to identify antibiotic resistance, but it takes 24-48 hours to obtain test results [6,7]. However, in the clinical workflow, clinicians need to identify antibiotic resistance quickly to provide effective treatment for patients with UTIs. For this reason, early prediction of antibiotic resistance in patients with UTIs is important to guide the selection of appropriate antibiotic therapy. Machine learning can be used to develop prediction models and clinical decision support systems (CDSSs) to identify antibiotic resistance and support the selection of appropriate antibiotic therapy for patients with a UTI.

Several efforts have been made to predict antibiotic resistance in patients with UTIs using data from patients' electronic medical records (EMRs), including demographics, prescriptions, comorbidities, procedures, and laboratory tests. These investigations have yielded promising results. Some of these studies were limited to specific patient groups, including patients with uncomplicated UTIs [8] and patients treated in the emergency department [9]. In other studies, researchers worked with heterogeneous data that were not limited to a specific patient group [10-12]. However, prior studies that analyzed heterogeneous data did not address the interpretation of machine

learning models. The black-box nature of machine learning is a limiting factor not only in its use for antibiotic resistance prediction but also in its wider clinical use [13,14]. Thus, interpreting the results obtained by the machine learning model is crucial in increasing users' trust in the machine learning model [15,16]. Furthermore, these studies did not address the development of the CDSS with the prediction models they built.

In this study, we aimed to predict antibiotic resistance in patients with a UTI. Heterogeneous data that were not limited to a specific patient group were used. UTI pathogens were classified as either sensitive or resistant to 5 commonly used antibiotics in UTI treatment: cephalosporin, piperacillin-tazobactam (TZP), carbapenem, trimethoprim-sulfamethoxazole (TMP-SMX), and fluoroquinolone. In addition, our objective was to understand and explain the inner workings of the machine learning models we developed. Eventually, a prototype CDSS was developed to provide clinicians the resistance probabilities for each antibiotic.

Methods

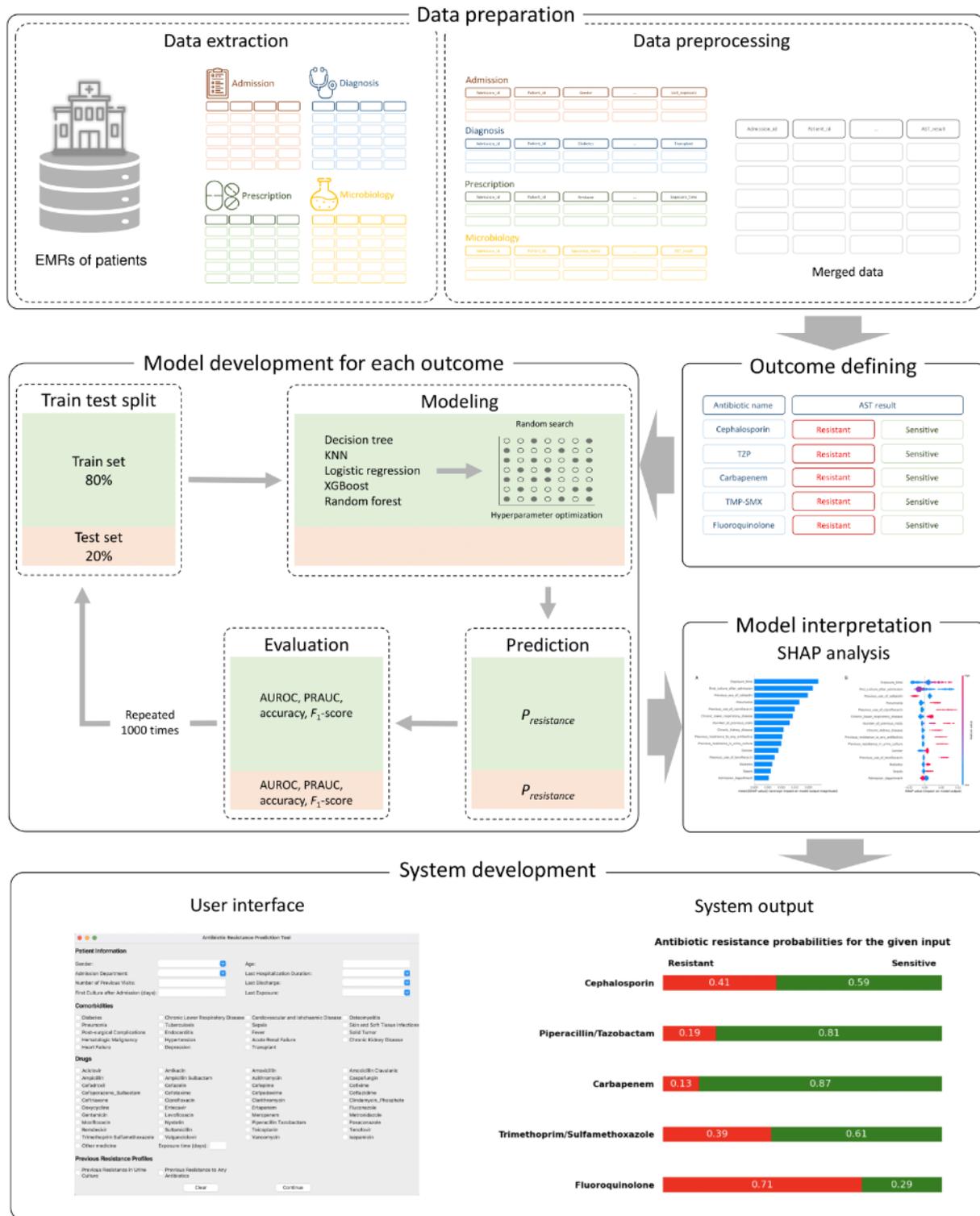
Ethical Considerations

Ethics approval for the study was obtained from the institutional review board of Yonsei University Severance Hospital on June 6, 2022 (approval 9-2023-0095). The informed consent was not required due to the retrospective nature of the study.

Data Set Description and Study Design

In this study, we used the EMRs of patients who were admitted to Yongin Severance Hospital, South Korea, between October 2012 and October 2022. To build the prediction models, admission, diagnosis, prescription, and microbiology records were extracted. The summary of the research process is presented in [Figure 1](#).

Figure 1. Summary of the research process. AST: antibiotic susceptibility test; AUROC: area under the receiver operating characteristic curve; EMR: electronic medical record; KNN: k-nearest neighbor; PRAUC: precision-recall area under the curve; SHAP: Shapley Additive Explanations; TMP-SMX: trimethoprim-sulfamethoxazole; TZP: piperacillin-tazobactam; XGBoost: Extreme Gradient Boosting.



Data Preprocessing

The microbiology table contained 143,114 urine cultures collected from 6011 patients during 7719 admissions. Since positive samples typically indicate the presence of bacteriuria, and urine culture samples were typically collected from patients with UTI symptoms, we considered these to be indicative of a UTI [10]. The resistance profiles were evaluated based on the

Clinical and Laboratory Standards Institute guidelines, where intermediate-level resistance was considered sensitive. To assess the resistance of UTI pathogens to antibiotic classes, antibiotics were grouped as cephalosporin, TZP, carbapenem, TMP-SMX, and fluoroquinolone. The antibiotics included in each antibiotic class are presented in [Multimedia Appendix 1](#). The patients' demographic information was extracted from the admission

table, their comorbidities were extracted from the diagnosis table, and their drug use information was extracted from the prescription table. For all input variables, the time of the first culture test was considered as the end point, and only data collected before the first culture test were used. After preprocessing and variable extraction from the raw data, the tables were combined using the admission number as the primary key. Missing data were excluded from the study. Patients aged 19 years and older and 100 years and younger at admission were included in the study, and numerical variables were standardized. A total of 71 features were used to classify UTI pathogens as either sensitive or resistant to each antibiotic. The predictors for the prediction models were selected by considering related works and using clinical judgment. Additionally, the threshold values for binarization were selected according to the literature [17] and the expert assessment of a specialist in infectious diseases. Detailed information about the predictors can be found in [Multimedia Appendix 2](#).

Machine Learning Model Development

We used a repeated train test split approach for modeling. The data sets were split into training and test sets using an 80:20 ratio, and the training sets were used for the development of the machine learning models. When splitting the data into training and test sets, data points from the same patient and admission were exclusively included in either the training or test data set to prevent potential data leakage and ensure the models were evaluated on previously unseen data. At each iteration, we created different training and test data sets by changing the random seed. Decision tree, k-nearest neighbor, logistic regression, Extreme Gradient Boosting, and random forest were used for modeling. The hyperparameters of the machine learning models were optimized by using the random search hyperparameter optimization method with 10-fold cross-validation on the training data set. We stored the performance of the prediction models at each iteration, and the mean of performance metrics was calculated. The procedure of splitting the data, optimizing hyperparameters, modeling, and evaluation was iteratively repeated 1000 times to classify UTI pathogens as either sensitive or resistant to cephalosporin, TZP, carbapenem, TMP-SMX, and fluoroquinolone. The machine learning models were built using Python (version 3.10.4; Python Software Foundation).

Machine Learning Model Interpretation

To analyze the contribution of the variables to the machine learning models in predicting antibiotic resistance, we used the Shapley Additive Explanations (SHAP) method. The SHAP values of the random forest models that showed superior performance compared to other machine learning methods were evaluated. The random forest model with the highest area under the receiver operating characteristic curve (AUROC) on the test

set across all iterations for each antibiotic was used for SHAP analysis. Python (version 3.10.4; Python Software Foundation) was used for SHAP analysis.

CDSS Development

To develop the CDSS prototype, the random forest model with the highest AUROC on the test set across all iterations for each antibiotic was used. The CDSS prototype was developed using the *tkinter* package in Python (version 3.10.4; Python Software Foundation).

Evaluation

The performance of the machine learning model for predicting antibiotic resistance was evaluated on the training and test sets using the AUROC with 95% CIs, precision-recall area under the curve (PRAUC), accuracy, and F_1 -score performance metrics. Herein, the AUROC value was considered the main evaluation metric. The definitions of the performance metrics we used are provided below.

- **AUROC:** The AUROC is a widely used metric that represents a classifier's ability to discriminate between positive instances and negative instances [18].
- **PRAUC:** PRAUC refers to the area under the precision-recall curve that plots precision as a function of recall for all the possible decision thresholds [19].
- **Accuracy:** Accuracy is the ratio of correctly classified samples to all samples.

$$\text{Accuracy} = \frac{\text{True positive} + \text{true negative}}{\text{True positive} + \text{false positive} + \text{false negative} + \text{true negative}}$$

- **F_1 -score:** F_1 -score is the harmonic mean of precision and recall metrics.

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{false positive}}$$

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{false negative}}$$

$$F_1\text{-score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{Precision} + \text{recall}}$$

Python (version 3.10.4; Python Software Foundation) was used to evaluate the prediction models.

Results

Data Set Characteristics

The general characteristics of the data set used in this study are presented in [Table 1](#). The data set included 3535, 737, 708, 1582, and 1365 samples for cephalosporin, TZP, TMP-SMX, fluoroquinolone, and carbapenem resistance prediction models, respectively. *Escherichia coli* was the most frequently isolated bacterial specimen across all antibiotics.

Table 1. General characteristics of the data set.

	Cephalosporin	TZP ^a	TMP-SMX ^b	Fluoroquinolone	Carbapenem
Samples, n	3535	737	708	1582	1365
Admissions, n	396	366	374	571	392
Patients, n	390	360	368	557	386
Resistance, n (%)	1492 (42.2)	169 (22.9)	281 (39.7)	1014 (64.1)	142 (10.4)
Age (years), mean (SD)	71.5 (14.4)	71.4 (14.4)	71.4 (14.4)	71.9 (14.4)	71.7 (14.3)
Female, n (%)	2597 (73.5)	523 (71)	507 (71.6)	1013 (64)	994 (72.8)
Most common bacteria (<i>Escherichia coli</i>), n (%)	1650 (46.7)	312 (42.3)	331 (46.7)	349 (22)	624 (45.7)
Second-most common bacteria (<i>Klebsiella pneumoniae</i>), n (%)	556 (15.7)	109 (14.8)	111 (15.7)	305 (19.3) ^c	220 (16.1)
Third-most common bacteria (<i>Pseudomonas aeruginosa</i>), n (%)	168 (4.7)	69 (9.4)	21 (3) ^d	180 (11.4) ^e	83 (6.1)

^aTZP: piperacillin-tazobactam.

^bTMP-SMX: trimethoprim-sulfamethoxazole.

^cThe isolated bacterial specimen is *Enterococcus faecium*.

^dThe isolated bacterial specimen is *Citrobacter freundii*.

^eThe isolated bacterial specimen is *Enterococcus faecalis*.

Model Performance

The performance analysis of the random forest models is presented in [Table 2](#). The AUROC values were 0.777 (95% CI 0.775-0.779), 0.864 (95% CI 0.862-0.867), 0.877 (95% CI 0.874-0.880), 0.881 (95% CI 0.879-0.882), and 0.884 (95% CI 0.884-0.885) in the training set and 0.638 (95% CI 0.635-0.642),

0.630 (95% CI 0.626-0.634), 0.665 (95% CI 0.659-0.671), 0.670 (95% CI 0.666-0.673), and 0.721 (95% CI 0.718-0.724) in the test set for predicting resistance to cephalosporin, TZP, carbapenem, TMP-SMX, and fluoroquinolone, respectively. The performance analysis of the other machine learning models is presented in [Multimedia Appendices 3-6](#).

Table 2. Classification performances of the random forest models.

	Training set				Test set			
	AUROC ^a (95% CI)	PRAUC ^b	Accuracy	F_1 -score	AUROC (95% CI)	PRAUC	Accuracy	F_1 -score
Cephalosporin	0.777 (0.775-0.779)	0.725	0.715	0.676	0.638 (0.635-0.642)	0.547	0.603	0.556
TZP ^c	0.864 (0.862-0.867)	0.688	0.808	0.652	0.630 (0.626-0.634)	0.332	0.641	0.313
Carbapenem	0.877 (0.874-0.880)	0.539	0.822	0.493	0.665 (0.659-0.671)	0.222	0.725	0.220
TMP-SMX ^d	0.881 (0.879-0.882)	0.829	0.822	0.781	0.670 (0.666-0.673)	0.568	0.638	0.560
Fluoroquinolone	0.884 (0.884-0.885)	0.938	0.802	0.832	0.721 (0.718-0.724)	0.813	0.657	0.706

^aAUROC: area under the receiver operating characteristic curve.

^bPRAUC: precision-recall area under the curve.

^cTZP: piperacillin-tazobactam.

^dTMP-SMX: trimethoprim-sulfamethoxazole.

Important Features

The SHAP values of the 15 most important features in the random forest models are presented in [Figure 2](#).

The SHAP feature importance bar plot ([Figure 3A](#)) and SHAP summary plot ([Figure 3B](#)) of the fluoroquinolone resistance prediction model are presented in [Figure 3](#). The SHAP feature importance plot and SHAP summary plot of the other antibiotic prediction models are presented in [Multimedia Appendices 7-10](#).

Figure 2. SHAP values of the 15 most important features in the prediction models. SHAP: Shapley Additive Explanations; TMP-SMX: trimethoprim-sulfamethoxazole; TZP: piperacillin-tazobactam.

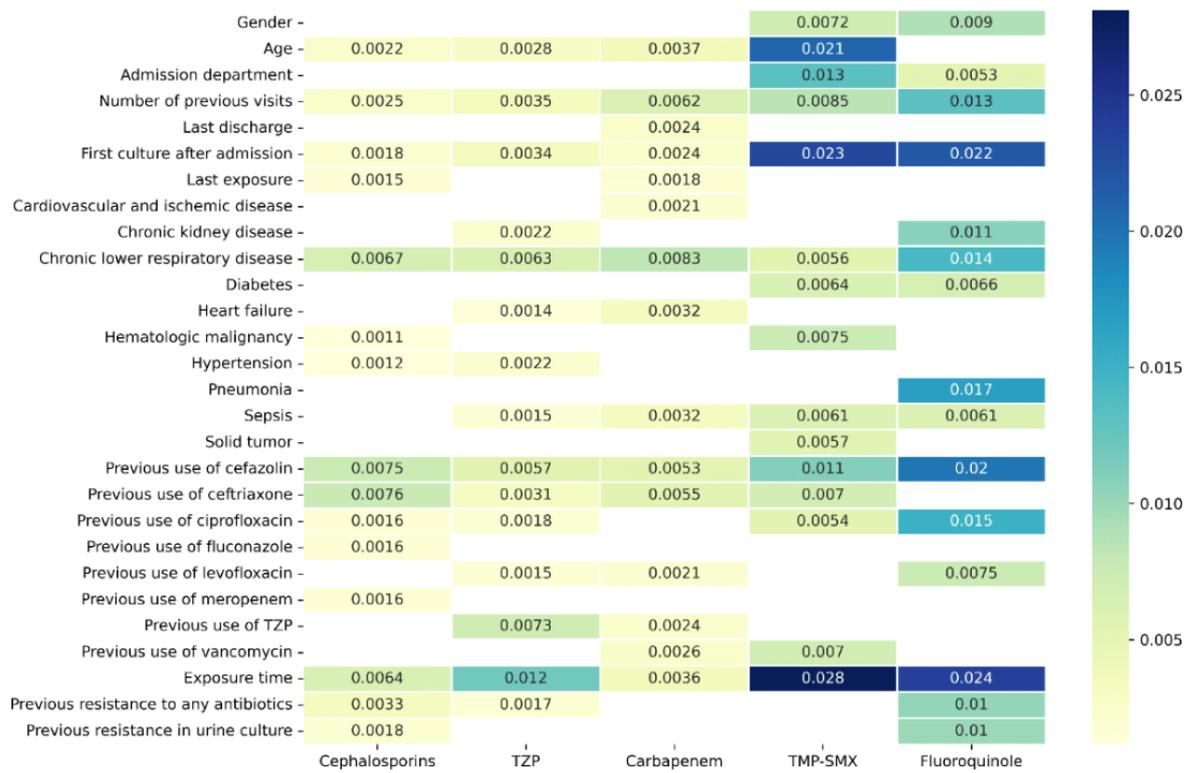
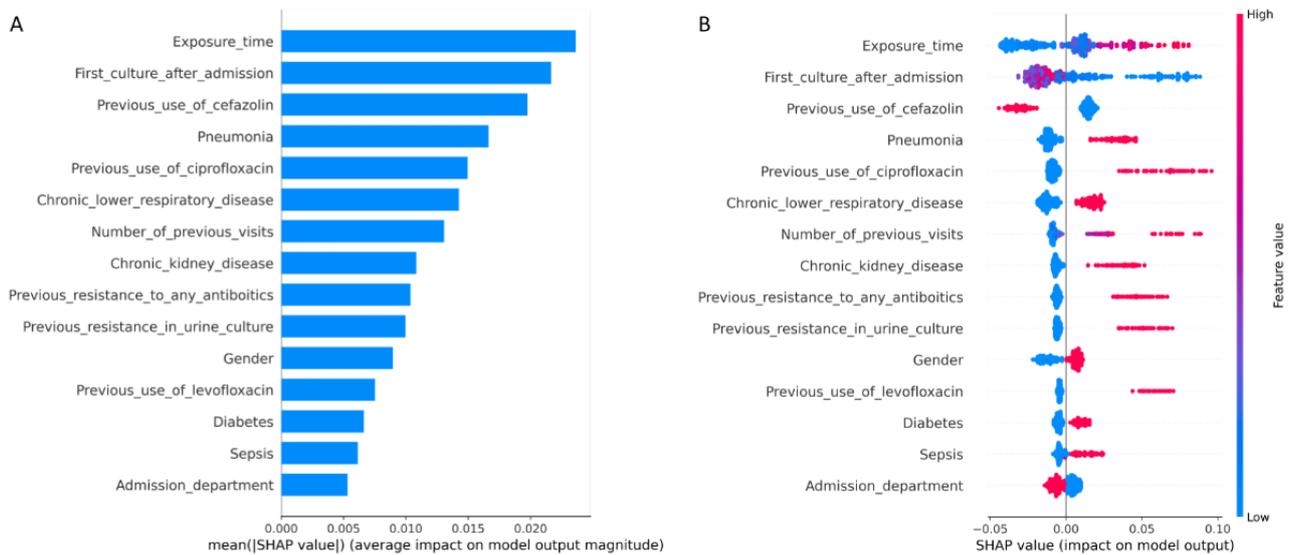


Figure 3. SHAP analysis results of fluoroquinolone resistance prediction model. (A) The feature importance bar plot. (B) The SHAP summary dot plot. SHAP: Shapley Additive Explanations.



Clinical Decision Support System

The user interface of the CDSS is shown in Figure 4. The CDSS prototype obtains data from the user and produces antibiotic resistance probabilities for each antibiotic.

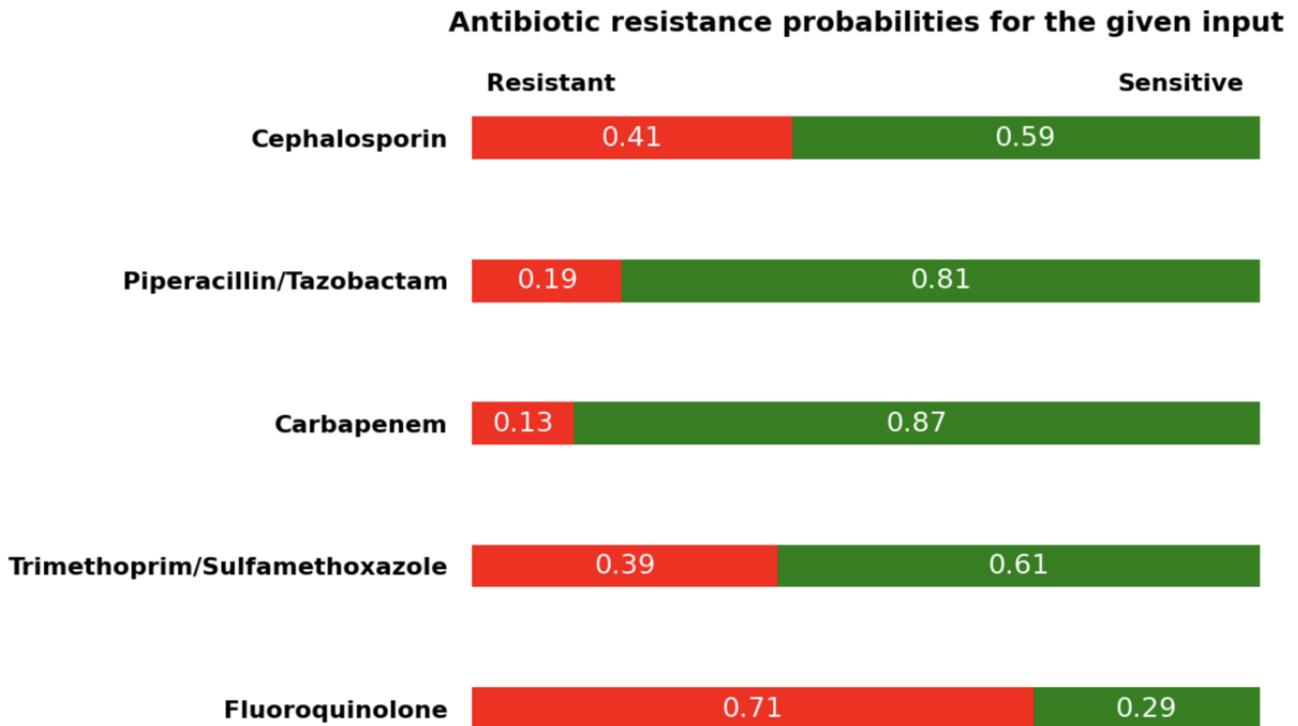
We presented the CDSS prototype on a scenario. In this case, a female aged 55 years was admitted to the hospital’s outpatient department. The patient previously visited the hospital 3 times and was readmitted to the hospital within 30 days of her last 3-day stay. The duration between the patient’s admission to the hospital and the first culture was 1 day. The patient was

previously diagnosed with diabetes and chronic lower respiratory disease. Additionally, the patient had a history of cefazolin use in the last 30 days and resistance in urine culture.

The system output for the given scenario is shown in Figure 5. The system produced resistance probabilities for each antibiotic. For the given scenario, the system produced a 71% probability of fluoroquinolone resistance, a 41% probability of cephalosporin resistance, a 39% probability of TMP-SMX resistance, a 19% probability of TZP resistance, and a 13% probability of carbapenem resistance.

Figure 4. The user interface of the clinical decision support system.

Figure 5. The screenshot of system output for the given data.



Discussion

Principal Findings

In this study, our main objective was to predict cephalosporin, TZP, carbapenem, TMP-SMX, and fluoroquinolone resistance in patients with UTI and develop a CDSS with the machine learning models we built. Moreover, we identified the most important features for predicting antibiotic resistance in patients with UTI using SHAP analysis.

Our prediction models achieved AUROCs of 0.777 (95% CI 0.775-0.779), 0.864 (95% CI 0.862-0.867), 0.877 (95% CI 0.874-0.880), 0.881 (95% CI 0.879-0.882), and 0.884 (95% CI 0.884-0.885) in the training set and 0.638 (95% CI 0.635-0.642), 0.630 (95% CI 0.626-0.634), 0.665 (95% CI 0.659-0.671), 0.670 (95% CI 0.666-0.673), and 0.721 (95% CI 0.718-0.724) in the test set for predicting resistance to cephalosporin, TZP, carbapenem, TMP-SMX, and fluoroquinolone, respectively. The fluoroquinolone resistance prediction model showed superior performance, as confirmed by its high AUROC values in both the training and test sets. On the other hand, the cephalosporin resistance prediction model showed poor performance, as confirmed by the low AUROC values in both training and test sets.

According to SHAP analysis, the contribution of the variables varied for each antibiotic; however, we found that the number of previous visits, first culture after admission, chronic lower respiratory diseases, administration of drugs before infection, and exposure time to these drugs were important predictors across all antibiotics. Factors such as the first culture after admission, exposure time, and the number of previous visits were found to affect resistance, which can be explained by the impact of health care-associated infections. Chronic lower respiratory and kidney diseases are also likely to be associated with frequent visits to health care facilities, although it is difficult to confirm the actual number of visits. However, this suggests that the characteristics of health care-seeking behavior in patients with specific underlying diseases may influence resistance [20]. Interestingly, the use of cefazolin had a negative impact on the development of resistance for all antibiotics. This is because cefazolin is one of the narrow-spectrum antibiotics used in less severe patients. Further research is needed to examine these results.

Comparison to Prior Work

Past efforts to predict antibiotic resistance in patients with UTIs have had promising results, with the lowest AUROC being 0.58 for predicting TMP-SMX resistance [12] and the highest AUROC being 0.83 for predicting ciprofloxacin resistance [9]. In comparison, our prediction models demonstrated comparable performance to these prior works. Some previous studies on predicting antibiotic resistance in patients with UTIs were limited to specific patient groups, including patients with uncomplicated UTIs [8] and patients treated in the emergency

department [9]. We analyzed heterogeneous data that were not limited to a specific patient group or bacteria. This approach provides a more comprehensive insight into the prediction of antibiotic resistance in patients with UTIs. Similarly, Lewin-Epstein et al [21] analyzed heterogeneous data and were able to achieve AUROC values ranging from 0.73 to 0.79 for the prediction of ceftazidime, gentamicin, imipenem, ofloxacin, and TMP-SMX resistance. Their data contained multiple culture tests, which provided a more comprehensive approach to predicting antibiotic resistance. Although urine cultures can be used to infer colonized resistance in patients, further research is needed to extend culture results beyond urine.

Limitations

While this study provides insights into predicting antibiotic resistance in patients with UTIs, it has some limitations. First, this study is the lack of multidrug resistance classification. The data set we used in this study did not contain a sufficient amount of multidrug resistance outcomes to build a classification model for the prediction of multidrug resistance. Furthermore, our prediction models were developed using prescription records within the hospital setting. However, patients may have used antibiotics outside of the hospital setting during visits to other hospitals. The lack of information about past drug use could have negatively impacted the performance of our prediction models. To overcome this limitation, we intend to conduct further studies using data from the National Health Insurance Service of South Korea, which contain all past drug use information of the patients. Thus, we will have a more comprehensive data set. By using this approach, we may be able to develop more accurate machine learning models to predict antibiotic resistance and improve our ability to guide appropriate antibiotic therapy selection. Additionally, further development is required to address the limitations of prototype CDSS, including the integration of real-time patient data and validation in larger patient cohorts. Moreover, the prototype CDSS only gives the resistance risk probability to the user. However, a more comprehensive system that can provide decision support on the selection of appropriate therapy, dosage, and duration of treatment can be developed in further studies. Such a system has the potential to reduce the duration of treatment, number of antibiotics used, cost, mortality, and morbidity [22,23].

Conclusions

In conclusion, our study results demonstrated that prediction models to predict antibiotic resistance in patients with UTIs can be constructed using routinely collected EMR data alone, without requiring additional laboratory tests or specialized tests. Machine learning techniques can be used to develop systems that can guide clinicians in selecting appropriate antibiotic therapy. This has the potential to prevent the risk of inappropriate antibiotic administration, thereby reducing patients' risk of developing antibiotic resistance.

Acknowledgments

This study was supported by the National Institute for International Education of the Government of the Republic of Korea, The Scientific and Technological Research Council of Turkey (grant 2214-A), and a faculty research grant of Yonsei University College of Medicine (6-2022-0118).

Authors' Contributions

NI, AY, SYP, and DY contributed to the conceptualization of the study and to the funding acquisition. JK, JAR, and SYP were responsible for data curation. NI performed the formal analysis of the collected data and wrote the paper. NI, SYP, and DY contributed to the development of the study methodology. SYP and DY reviewed and edited the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The antibiotics included in each antibiotic class.

[\[DOCX File , 17 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Description of input variables.

[\[DOCX File , 15 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Classification performances of the decision tree models.

[\[DOCX File , 20 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Classification performances of the k-nearest neighbor models.

[\[DOCX File , 20 KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

Classification performances of the logistic regression models.

[\[DOCX File , 20 KB-Multimedia Appendix 5\]](#)

Multimedia Appendix 6

Classification performances of the Extreme Gradient Boosting models.

[\[DOCX File , 20 KB-Multimedia Appendix 6\]](#)

Multimedia Appendix 7

SHAP analysis results of cephalosporin resistance prediction model. (A) The feature importance bar plot. (B) The SHAP summary dot plot. SHAP: Shapley Additive Explanations.

[\[PNG File , 121 KB-Multimedia Appendix 7\]](#)

Multimedia Appendix 8

SHAP analysis results of TZP resistance prediction model. (A) The feature importance bar plot. (B) The SHAP summary dot plot. SHAP: Shapley Additive Explanations; TZP: piperacillin-tazobactam.

[\[PNG File , 117 KB-Multimedia Appendix 8\]](#)

Multimedia Appendix 9

SHAP analysis results of carbapenem resistance prediction model. (A) The feature importance bar plot. (B) The SHAP summary dot plot. SHAP: Shapley Additive Explanations; TZP: piperacillin-tazobactam.

[\[PNG File , 122 KB-Multimedia Appendix 9\]](#)

Multimedia Appendix 10

SHAP analysis results of TMP-SMX resistance prediction model. (A) The feature importance bar plot. (B) The SHAP summary dot plot. SHAP: Shapley Additive Explanations; TMP-SMX: trimethoprim-sulfamethoxazole.

[\[PNG File , 123 KB-Multimedia Appendix 10\]](#)

References

1. Tan CW, Chlebicki MP. Urinary tract infections in adults. *Singapore Med J.* 2016;57(9):485-490. [\[FREE Full text\]](#) [doi: [10.11622/smedj.2016153](https://doi.org/10.11622/smedj.2016153)] [Medline: [27662890](https://pubmed.ncbi.nlm.nih.gov/27662890/)]
2. Belete MA, Saravanan M. A systematic review on drug resistant urinary tract infection among pregnant women in developing countries in Africa and Asia; 2005-2016. *Infect Drug Resist.* 2020;13:1465-1477. [\[FREE Full text\]](#) [doi: [10.2147/IDR.S250654](https://doi.org/10.2147/IDR.S250654)] [Medline: [32547115](https://pubmed.ncbi.nlm.nih.gov/32547115/)]
3. Santos M, Mariz M, Tiago I, Martins J, Alarico S, Ferreira P. A review on urinary tract infections diagnostic methods: laboratory-based and point-of-care approaches. *J Pharm Biomed Anal.* 2022;219:114889. [\[FREE Full text\]](#) [doi: [10.1016/j.jpba.2022.114889](https://doi.org/10.1016/j.jpba.2022.114889)] [Medline: [35724611](https://pubmed.ncbi.nlm.nih.gov/35724611/)]
4. Suskind AM, Saigal CS, Hanley JM, Lai J, Setodji CM, Clemens JQ, et al. Urologic Diseases of America Project. Incidence and management of uncomplicated recurrent urinary tract infections in a national sample of women in the United States. *Urology.* 2016;90:50-55. [\[FREE Full text\]](#) [doi: [10.1016/j.urology.2015.11.051](https://doi.org/10.1016/j.urology.2015.11.051)] [Medline: [26825489](https://pubmed.ncbi.nlm.nih.gov/26825489/)]
5. Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol.* 2015;13(5):269-284. [\[FREE Full text\]](#) [doi: [10.1038/nrmicro3432](https://doi.org/10.1038/nrmicro3432)] [Medline: [25853778](https://pubmed.ncbi.nlm.nih.gov/25853778/)]
6. Boolchandani M, D'Souza AW, Dantas G. Sequencing-based methods and resources to study antimicrobial resistance. *Nat Rev Genet.* 2019;20(6):356-370. [\[FREE Full text\]](#) [doi: [10.1038/s41576-019-0108-4](https://doi.org/10.1038/s41576-019-0108-4)] [Medline: [30886350](https://pubmed.ncbi.nlm.nih.gov/30886350/)]
7. Benkova M, Soukup O, Marek J. Antimicrobial susceptibility testing: currently used methods and devices and the near future in clinical practice. *J Appl Microbiol.* 2020;129(4):806-822. [doi: [10.1111/jam.14704](https://doi.org/10.1111/jam.14704)] [Medline: [32418295](https://pubmed.ncbi.nlm.nih.gov/32418295/)]
8. Kanjilal S, Oberst M, Boominathan S, Zhou H, Hooper DC, Sontag D. A decision algorithm to promote outpatient antimicrobial stewardship for uncomplicated urinary tract infection. *Sci Transl Med.* 2020;12(568):eaay5067. [\[FREE Full text\]](#) [doi: [10.1126/scitranslmed.aay5067](https://doi.org/10.1126/scitranslmed.aay5067)] [Medline: [33148625](https://pubmed.ncbi.nlm.nih.gov/33148625/)]
9. Lee HG, Seo Y, Kim JH, Han SB, Im JH, Jung CY, et al. Machine learning model for predicting ciprofloxacin resistance and presence of ESBL in patients with UTI in the ED. *Sci Rep.* 2023;13(1):3282. [\[FREE Full text\]](#) [doi: [10.1038/s41598-023-30290-y](https://doi.org/10.1038/s41598-023-30290-y)] [Medline: [36841917](https://pubmed.ncbi.nlm.nih.gov/36841917/)]
10. Yelin I, Snitser O, Novich G, Katz R, Tal O, Parizade M, et al. Personal clinical history predicts antibiotic resistance of urinary tract infections. *Nat Med.* 2019;25(7):1143-1152. [\[FREE Full text\]](#) [doi: [10.1038/s41591-019-0503-6](https://doi.org/10.1038/s41591-019-0503-6)] [Medline: [31273328](https://pubmed.ncbi.nlm.nih.gov/31273328/)]
11. Hebert C, Gao Y, Rahman P, Dewart C, Lustberg M, Pancholi P, et al. Prediction of antibiotic susceptibility for urinary tract infection in a hospital setting. *Antimicrob Agents Chemother.* 2020;64(7):e02236-19. [\[FREE Full text\]](#) [doi: [10.1128/AAC.02236-19](https://doi.org/10.1128/AAC.02236-19)] [Medline: [32312778](https://pubmed.ncbi.nlm.nih.gov/32312778/)]
12. Rich SN, Jun I, Bian J, Boucher C, Cherabuddi K, Morris JG, et al. Development of a prediction model for antibiotic-resistant urinary tract infections using integrated electronic health records from multiple clinics in North-Central Florida. *Infect Dis Ther.* 2022;11(5):1869-1882. [\[FREE Full text\]](#) [doi: [10.1007/s40121-022-00677-x](https://doi.org/10.1007/s40121-022-00677-x)] [Medline: [35908268](https://pubmed.ncbi.nlm.nih.gov/35908268/)]
13. Watson DS, Krutzinna J, Bruce IN, Griffiths CE, McInnes IB, Barnes MR, et al. Clinical applications of machine learning algorithms: beyond the black box. *BMJ.* 2019;364:1886. [doi: [10.1136/bmj.1886](https://doi.org/10.1136/bmj.1886)] [Medline: [30862612](https://pubmed.ncbi.nlm.nih.gov/30862612/)]
14. Macesic N, Polubriaginof F, Tatonetti NP. Machine learning: novel bioinformatics approaches for combating antimicrobial resistance. *Curr Opin Infect Dis.* 2017;30(6):511-517. [doi: [10.1097/QCO.0000000000000406](https://doi.org/10.1097/QCO.0000000000000406)] [Medline: [28914640](https://pubmed.ncbi.nlm.nih.gov/28914640/)]
15. Tucci V, Saary J, Doyle TE. Factors influencing trust in medical artificial intelligence for healthcare professionals: a narrative review. *J Med Artif Intell.* 2022;5:4-4. [\[FREE Full text\]](#) [doi: [10.21037/jmai-21-25](https://doi.org/10.21037/jmai-21-25)]
16. Elshawi R, Al-Mallah MH, Sakr S. On the interpretability of machine learning-based model for predicting hypertension. *BMC Med Inform Decis Mak.* 2019;19(1):146. [\[FREE Full text\]](#) [doi: [10.1186/s12911-019-0874-0](https://doi.org/10.1186/s12911-019-0874-0)] [Medline: [31357998](https://pubmed.ncbi.nlm.nih.gov/31357998/)]
17. Tacconelli E. New strategies to identify patients harbouring antibiotic-resistant bacteria at hospital admission. *Clin Microbiol Infect.* 2006;12(2):102-109. [\[FREE Full text\]](#) [doi: [10.1111/j.1469-0691.2005.01326.x](https://doi.org/10.1111/j.1469-0691.2005.01326.x)] [Medline: [16441446](https://pubmed.ncbi.nlm.nih.gov/16441446/)]
18. Janssens ACJW, Martens FK. Reflection on modern methods: revisiting the area under the ROC curve. *Int J Epidemiol.* 2020;49(4):1397-1403. [\[FREE Full text\]](#) [doi: [10.1093/ije/dy274](https://doi.org/10.1093/ije/dy274)] [Medline: [31967640](https://pubmed.ncbi.nlm.nih.gov/31967640/)]
19. Cook J, Ramadas V. When to consult precision-recall curves. *Stata J.* 2020;20(1):131-148. [\[FREE Full text\]](#) [doi: [10.1177/1536867x20909693](https://doi.org/10.1177/1536867x20909693)]
20. Park H, Son MJ, Jung DW, Lee H, Lee JY. National trends in hospitalization for ambulatory care sensitive conditions among Korean adults between 2008 and 2019. *Yonsei Med J.* 2022;63(10):948-955. [\[FREE Full text\]](#) [doi: [10.3349/ymj.2022.0110](https://doi.org/10.3349/ymj.2022.0110)] [Medline: [36168248](https://pubmed.ncbi.nlm.nih.gov/36168248/)]

21. Lewin-Epstein O, Baruch S, Hadany L, Stein GY, Obolski U. Predicting antibiotic resistance in hospitalized patients by applying machine learning to electronic medical records. *Clin Infect Dis*. 2021;72(11):e848-e855. [FREE Full text] [doi: [10.1093/cid/ciaa1576](https://doi.org/10.1093/cid/ciaa1576)] [Medline: [33070171](https://pubmed.ncbi.nlm.nih.gov/33070171/)]
22. Curtis CE, Al Bahar F, Marriott JF. The effectiveness of computerised decision support on antibiotic use in hospitals: a systematic review. *PLoS One*. 2017;12(8):e0183062. [FREE Full text] [doi: [10.1371/journal.pone.0183062](https://doi.org/10.1371/journal.pone.0183062)] [Medline: [28837665](https://pubmed.ncbi.nlm.nih.gov/28837665/)]
23. Laka M, Milazzo A, Merlin T. Can evidence-based decision support tools transform antibiotic management? A systematic review and meta-analyses. *J Antimicrob Chemother*. 2020;75(5):1099-1111. [FREE Full text] [doi: [10.1093/jac/dkz543](https://doi.org/10.1093/jac/dkz543)] [Medline: [31960021](https://pubmed.ncbi.nlm.nih.gov/31960021/)]

Abbreviations

AUROC: area under the receiver operating characteristic curve

CDSS: clinical decision support system

EMR: electronic medical record

PRAUC: precision-recall area under the curve

SHAP: Shapley Additive Explanations

TMP-SMX: trimethoprim-sulfamethoxazole

TZP: piperacillin-tazobactam

UTI: urinary tract infection

Edited by A Benis; submitted 27.07.23; peer-reviewed by MO Khurshed, YJ Tseng; comments to author 25.08.23; revised version received 17.11.23; accepted 08.01.24; published 29.02.24

Please cite as:

İlhanlı N, Park SY, Kim J, Ryu JA, Yardımcı A, Yoon D

Prediction of Antibiotic Resistance in Patients With a Urinary Tract Infection: Algorithm Development and Validation

JMIR Med Inform 2024;12:e51326

URL: <https://medinform.jmir.org/2024/1/e51326>

doi: [10.2196/51326](https://doi.org/10.2196/51326)

PMID: [38421718](https://pubmed.ncbi.nlm.nih.gov/38421718/)

©Nevruz İlhanlı, Se Yoon Park, Jaewoong Kim, Jee An Ryu, Ahmet Yardımcı, Dukyong Yoon. Originally published in JMIR Medical Informatics (<https://medinform.jmir.org>), 29.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.