Predicting Antimicrobial Resistance in Uncomplicated Urinary Tract Infections Using Machine-Learning

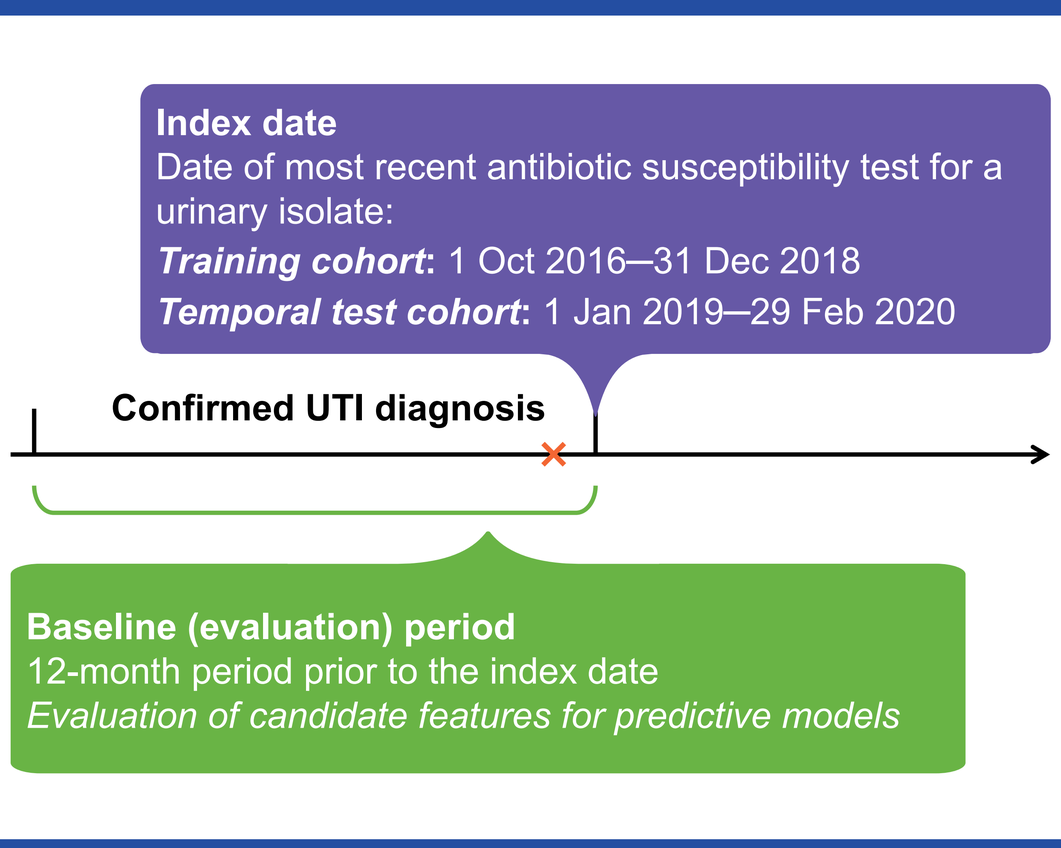
Kalé Kponee-Shovein,[1] Wendy Y Cheng,[1] Alen Marijam,[2] Patrick Schwab,[3] Chi Gao,[1] Daniel Indacochea,[1] Diogo Ferrinho,[3] Fanny S Mitrani-Gold,[3] Lisa Pinheiro,[1] Jimmy Royer,[1] Ashish V Joshi[3]

[1]Analysis Group, Inc., Boston, MA, USA; [2]GSK, Wavre, Belgium; [3]GSK, Collegeville, PA, USA

Introduction and Objectives

- Urinary tract infections (UTIs) are among the most common bacterial infections in the US, with approximately 80% classified as uncomplicated (uUTIs)[1]
- Over 50% of patients with uUTI are prescribed non-guideline-based antimicrobial treatment,[2] potentially contributing to antimicrobial resistance (AMR) and increased healthcare costs
- In the US, common first-line antibiotics prescribed for uUTI include nitrofurantoin (NFT) and trimethoprim/sulfamethoxazole (SXT), while second-line and other antibiotics include fluoroquinolones (FQs) and beta-lactams (BLs)[3]
- Predictive models using machine-learning techniques based on uUTI patient features could help clinicians identify individuals with a higher likelihood of AMR to specific classes of antibiotics
- Empiric prescribing can mean that antibiotics less likely to be effective are administered; machine-learning might inform options early in the treatment course of uUTI to lessen the health and economic burden of uUTI
- This study aimed to use machine-learning to develop a predictive algorithm for the identification of clinically relevant predictors of AMR to any of NFT, SXT, FQs, and BLs among female patients with uUTI

Figure. Study design



Methods

- In this retrospective observational study, demographic and clinical characteristics, including laboratory results, were obtained from the Optum Electronic Health Record (EHR) database (1 Oct 2015–29 Feb 2020)
- Female patients were eligible if they were ≥12 years of age with a uUTI diagnosis with pyuria confirmed by a positive *Escherichia coli* (*E. coli,* the most uropathogen in uUTI) urine culture within 7 days of diagnosis; patients were treated with ≥1 oral antibiotic of interest (NFT, SXT, FQ, or BL) and antibiotic susceptibility testing was performed for uropathogens within 7 days of diagnosis
- Patients were included in either the training cohort or the temporal test cohort according to index date (**Figure, Table 1**)
- Least Absolute Shrinkage and Selection Operator (LASSO) and Random Forest (RF) were evaluated as candidate predictive algorithms of AMR for each antibiotic treatment class in the training cohort

Methods (Continued)

- The optimal predictive algorithm was chosen based on the Area Under the Curve of the Receiver Operating Characteristic (AUROC) and clinical interpretability (**Table 2**)
- Multiple imputation by chained equations was used to impute missing data, and nested cross validation was used to select the optimal algorithm in the training cohort
- LASSO was deemed the optimal algorithm and its performance was then evaluated in the temporal test cohort (**Table 3**)

Table 1. The proportion of patients with confirmed antibiotic susceptibility test results of a urinary *E. coli* isolate, stratified into training and testing cohorts based on index date

Cohort Population Numbers, n (% of Drug Class)	NFT	SXT	FQ	BL
Training cohort (n=212,715)	**52,718** (64.1)	**53,221** (64.2)	**52,809** (65.0)	**53,967** (64.3)
Temporal test cohort (n=117,589)	**29,507** (35.9)	**29,700** (35.8)	**28,414** (35.0)	**29,968** (35.7)

Table 2. Nested cross validation AUROC values of LASSO and Random Forest algorithms for antibiotic treatment classes: Training cohort

Antibiotic Treatment Class	Mean AUROC Value* (SE)†	
	LASSO	Random Forest
NFT	**0.64** (0.021)	**0.63** (0.020)
SXT	**0.66** (0.009)	**0.65** (0.010)
FQ	**0.72** (0.012)	**0.71** (0.009)
BL	**0.66** (0.011)	**0.66** (0.012)

Table 3. AUROC values of LASSO algorithms for antibiotic treatment classes: Temporal test cohort

Antibiotic Treatment Class	Mean AUROC Value* (SE)‡
NFT	**0.67** (0.010)
SXT	**0.66** (0.004)
FQ	**0.72** (0.005)
BL	**0.66** (0.005)

SE, standard error. *The AUROC was generated by plotting the true positive rate versus the false positive rate resulting from different thresholds in the predictive model, and then calculating the area under the curve. †Standard errors were calculated across the 10 outer test folds during nested cross validation. ‡Standard errors were calculated by bootstrapping the test cohort and computing the corresponding metrics using the final prediction model, whose coefficients were estimated based on the full original training cohort.

Summary and Conclusions

- Machine-learning algorithms can be leveraged to aid healthcare professionals in the decision-making process; here, the algorithm will be developed into a tool to predict AMR risk for uropathogens in patients with uUTI, to inform correct empirical prescribing
- LASSO models were selected over RF models for the development and validation of predictive algorithms for all treatment classes due to better clinical interpretability
- The predictive algorithms developed for each antibiotic class provide good predictive power in distinguishing between patients with high versus low risk for uropathogens with AMR
- The Optum EHR dataset is larger and more generalisable than other databases used in published uUTI models in the US[4]; this predictive algorithm improves upon existing models by enabling greater statistical power and generalisability

References

1. Zilberberg M, et al. *Open Forum Infect Dis.* 2021;1427 (abstract).
2. Durkin MJ, et al. *Open Forum Infect Dis.* 2018;5:9.
3. Wagenlehner F, et al. *J Glob Antimicrob Resist.* 2022;28:18–29.
4. Kanjilal S, et al. *Sci Transl Med.* 2020;12:568.

Disclosures

AM, PS, DF, FSM-G, and AVJ are employees of, and shareholders in, GSK. KK-S, WYC, CG, DI, LP, and JR are employees of Analysis Group, Inc., which received funding from GSK to conduct the study. GSK-sponsored study 214849. Medical writing support for the development of this poster, under direction of the authors, was provided by Christian Jones, PhD, of Ashfield MedComms, an Inizio company, and funded by GSK.