

# The urinary microbiome distinguishes symptomatic urinary tract infection from asymptomatic older adult patients presenting to the emergency department

Evan S. Bradley, Celina Stansky , Abigail L. Zeamer, Ziyuan Huang, Lindsey Cincotta , Abigail Lopes , Linda Potter , Theresa Fontes , Doyle V. Ward, Vanni Bucci , Beth A. McCormick & John P. Haran

**To cite this article:** Evan S. Bradley, Celina Stansky , Abigail L. Zeamer, Ziyuan Huang, Lindsey Cincotta , Abigail Lopes , Linda Potter , Theresa Fontes , Doyle V. Ward, Vanni Bucci , Beth A. McCormick & John P. Haran (2025) The urinary microbiome distinguishes symptomatic urinary tract infection from asymptomatic older adult patients presenting to the emergency department, *Virulence*, 16:1, 2546063, DOI: [10.1080/21505594.2025.2546063](https://doi.org/10.1080/21505594.2025.2546063)

**To link to this article:** <https://doi.org/10.1080/21505594.2025.2546063>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 13 Aug 2025.



[Submit your article to this journal](#)



Article views: 225



[View related articles](#)



[View Crossmark data](#)

## The urinary microbiome distinguishes symptomatic urinary tract infection from asymptomatic older adult patients presenting to the emergency department

Evan S. Bradley , Celina Stansky<sup>a</sup>, Abigail L. Zeamer , Ziyuan Huang , Lindsey Cincotta<sup>a</sup>,  
Abigail Lopes<sup>a</sup>, Linda Potter<sup>d</sup>, Theresa Fontes<sup>d</sup>, Doyle V. Ward , Vanni Buccib<sup>c</sup>, Beth A. McCormick ,  
and John P. Haran 

<sup>a</sup>Department of Emergency Medicine, UMass Memorial Medical Center, Worcester, MA, USA; <sup>b</sup>Program in Microbiome Dynamics, UMass Chan Medical School, Worcester, MA, USA; <sup>c</sup>Department of Microbiology, UMass Chan Medical School, Worcester, MA, USA; <sup>d</sup>Clinical Renal Laboratory, UMass Memorial Medical Center, Worcester, MA, USA

### ABSTRACT

Older adults suffer from a high rate of asymptomatic bacteriuria (ASB), in which urinalysis may appear positive (presence of bacteria, white blood cells, and nitrates), often triggering initiation of antibiotics in acute care settings, without actual urinary tract infection (UTI) present. To investigate the urinary microbiome of older adults being tested for UTI, we enrolled a convenience sample of 250 older adult Emergency Department patients who had microscopic urinalysis ordered as part of their routine clinical care. Urinalysis results were classified as positive or negative, and patients were classified as being symptomatic or asymptomatic based on established diagnostic guidelines. We sought to determine if features of the urinary microbiome differed between positive and negative urinalysis (UAs) and symptomatic and asymptomatic patients with positive UAs. The same urine sample used for clinical testing was sequenced and analyzed for bacterial taxa, metabolic pathways, and known bacterial virulence factors. After exclusion of anatomical abnormalities and filtering for sequencing quality, 152 samples were analyzed (5 negative UAs, 147 positive UAs, among which 68 were asymptomatic, and 79 symptomatic). Positive UA samples showed significantly lower alpha diversity (2.29 versus 0.086,  $p < 0.01$ ) and distinct community composition based on beta-diversity (PERMANOVA on Bray-Curtis distance  $p < 0.01$ ). Alpha and beta diversity did not significantly differ between asymptomatic and symptomatic patients. Machine learning classifiers combining clinical covariates other than specific signs and symptoms and microbiome features (taxa, metabolic pathways, or virulence factors) revealed mostly microbiome features as predictive of symptomatic UTI over clinical features.

### ARTICLE HISTORY

Received 20 January 2025

Revised 8 July 2025

Accepted 4 August 2025

### KEYWORDS

Urinary tract infection;  
metagenomic analysis;  
machine learning; older  
adults; emergency  
department; asymptomatic  
bacteriuria

## Introduction

It was once thought that human urinary tracts were normally sterile environments, but it is now known that there is a community of microorganisms that resides in the urinary tract, known as the urinary microbiome or uromicrobiome [1–8]. Changes in the urinary microbiome have been connected to a wide variety of genitourinary conditions such as urinary urge incontinence, pelvic pain syndromes, and recurrent urinary tract infections (UTIs) [9,10]. Annually in the US, UTI is listed as the cause of some 1 million Emergency Department (ED) visits and 15% of acute care hospitalizations among patients 65 years of age and older [11]. UTI is a potentially life-threatening diagnosis as it can progress to sepsis and septic shock, which carries high mortality [12]. Despite its prevalence and impact on the healthcare system, appropriate diagnosis of UTIs

remains a challenge, especially in older adults. This is because a substantial portion of the population over age 65 have a condition known as asymptomatic bacteriuria (ASB) in which the urinary tract is colonized with potentially pathogenic bacteria, but the individual does not have symptoms and does not benefit from treatment with antibiotics [13]. ASB is defined by the Infectious Disease Society of America (IDSA) as a urine culture growing  $10^5$  colony forming units of a known urinary pathogen from an individual who is having no symptoms [14]. Since urine cultures can take up to 3 days, what is typically relied upon as a proxy in the ED are the results of microscopic urinalysis (UA) that are suggestive of infection, i.e. the presence of nitrates, the presence of leukocyte esterase greater than trace, or greater than 10 white blood cells per high-powered field on microscopy [15]. The IDSA only recommends

treating these patients with antibiotics only if they have signs and symptoms of a urinary tract infection such as dysuria, lower abdominal or perineal pain in the case of uncomplicated infection, or flank pain and fever in the case of complicated infection, coincident with UA findings that are suggestive of infection [14]. Guidance from the IDSA suggests that, except for pregnant women or those undergoing invasive urologic procedures, there is no benefit to treating ASB with antibiotics [16]. ASB prevalence is 10–20% in community-dwelling older adults [17] and as high as 50% in those who reside in skilled nursing facilities [13]. Among patients with positive urine cultures admitted to an inpatient service from the ED about 25% did not have signs or symptoms of UTI [18]. These discrepancies highlight the difficulties in accurately distinguishing ASB from UTI, especially in the acute care setting.

The clinical tests most frequently used to diagnose UTI, the UA and urine culture, have significant limitations. A study of adult ED patients suggested that a UA with features previously mentioned as associated with UTI had a specificity of 54% and a positive predictive value of 38% for a positive urine culture [15]. Because the UA alone lacks specificity for diagnosis of UTI, published guidelines suggest that UA should only be used in patients with signs and symptoms of UTI [19,20]. Despite this, clinicians frequently test for and initiate treatment for presumed UTI in the absence of signs or symptoms of infection [21]. A study of patients admitted from the ED to a general medical service found that as many as 80% of older patients had a UA ordered without the presence of urinary symptoms [22]. This discrepancy between actual practice and recommended guidelines may in part reflect the poor specificity of current standard diagnostic procedures as well as the desire of physicians to be thorough when presented with cases with multiple possible differential diagnoses, as is often the case in the ED setting.

The urinary microbiome may hold insights in distinguishing between ASB and UTI. There may be characteristics of the overall community that are suggestive of one state versus the other. An imbalance of pathogens over other commensal species in urine may be the microbiologic signature of shifting from ASB to UTI. Metagenomic data about the presence of specific virulence factor genes or metabolic pathways could provide additional clues. For example, Uropathogenic *Escherichia coli* (UPEC) typically require pathogenicity factors, often encoded on pathogenicity islands [23] to cause active disease; colonizing organisms in patients with ASB may lack genes for pathogenicity factors and thus, the presence or increased abundance of specific

genes may be a reliable signature of UTI. There are also likely metabolic requirements for organisms to establish infection in the urinary tract that could signal a transition to UTI. Given that the urinary microbiome field is relatively new and comprehensive characterization is emerging, more studies, especially in clinical populations, are warranted.

The primary objective of this study was to examine the urinary microbiome of older adult patients with suspected UTI on initial evaluation in the ED. Using metagenomic sequencing and machine learning-based multimodal analysis we sought to determine if urinary microbiome features showed reliable differences between symptomatic versus asymptomatic patients, with the goal of selecting highly discriminatory features that could potentially be investigated as diagnostic targets.

## Materials and methods

### Patient recruitment

We enrolled in a convenience sample of adult patients over 65 years old admitted to the UMass Memorial University Campus ED from January 2022 through April 2023 who had a urinalysis ordered and collected as part of routine clinical care. Written informed consent was obtained from all patients. If a patient was deemed not to have capacity by medical providers or documented in their chart, written informed consent was obtained from patient's legally authorized representative (next of kin or designated health care proxy). Those with potentially confounding anatomic abnormalities (e.g. urostomy or nephrostomy) and chronic indwelling Foley catheters on ED admission were excluded as these conditions may cause unknown long-term changes to the urinary microbiome not reflective of ASB. Patients who had a urinary catheter placed at the ED visit in which they were enrolled were not excluded, as this is frequently performed to obtain specimens from individuals who are incontinent of urine or otherwise cannot provide a sample on demand and represents an important population to study with respect to UTI. Patients were interviewed for current urinary symptoms including new onset dysuria (pain on urination), urinary frequency, urinary urgency, urinary incontinence, hematuria, urinary retention, lower abdominal pain, fever, or flank pain. Signs of urinary tract infection including lower abdominal tenderness, altered mental status, and costovertebral angle tenderness were obtained from electronic medical records. Patients were also asked about recent medical history including hospitalizations, surgeries, urinary

catheterization, recent antibiotic treatment, and recent urinary tract infections. Past medical history, current medications, ED disposition, final diagnosis were all obtained via chart review. A complete list of patient variables that were collected and used as covariates within the models is shown in Table S3. We also review study participants' medical charts up to 14 days post admission to determine if there were any additional unplanned hospitalizations, treatments for urinary tract infection, or death.

Urine samples were obtained by nursing staff in the ED either by mid-stream clean catch or urinary catheters in patients incontinent of urine or otherwise not able to provide a sample. Although as previously stated, anyone with a chronic indwelling urinary catheter on admission was excluded. Patients were categorized as having symptoms of UTI based on current diagnostic guidelines of signs and symptoms provided by IDSA and the Massachusetts Department of Public Health [13,19,20]. That is, the presence of dysuria considered most specific for UTI, or the presence of at least two signs or symptoms that could be attributed to UTI based on patient interview and physician chart review. These included fever, altered mental status, generalized weakness, lower abdominal pain, urinary frequency, urinary urgency, new urinary incontinence, new urinary retention, flank pain, or flank tenderness. Traditional urinalysis results were considered positive if there was at least one feature consistent with infection that is usually used by emergency physicians to make treatment decisions regarding UTI; these were the presence of nitrites, the presence of leukocyte esterase greater than trace, or greater than 10 white blood cells per high-powered field on microscopy [15]. The human subjects research protocol used for this study was approved by the UMass Chan Institutional Review Board (Docket # H00014312)

### **Sample collection and processing**

The remainder of the urine samples from enrolled patients provided to the clinical laboratory for traditional microscopic urinalysis and urine culture were kept at 4°C in the UMass Memorial Renal Laboratory until they were collected by study staff and then stored long term at -80°C. Samples were thawed and batch processed for sequencing in sets of 20 samples at a time. A total volume of 5 mL of each urine sample was used for metagenomic analysis. Left-over samples of less than 5 mL of urine were diluted with DNA/RNA shield solution (Zymo Research) to 5 mL each. Eukaryotic cell contaminants were first removed by low-speed spin at 180 g ×15 min at 4°C. Supernatants

were then pelleted for bacteria by spinning at 3000 rcf ×15 min and washed once with 5 mL of DNA/RNA shield solution. The final pellet was resuspended in 250 µL of DNA/RNA shield and total DNA extracted using the Zymo Research DNA miniprep kit with bead-beating step. Samples were prepared for paired-end sequencing using the Nextera XT DNA library preparation kit (Illumina) and sequencing was performed on the Illumina NextSeq at the UMass Chan Center for Microbiome Research. Negative controls were not included in library preparation and sequencing and corrections for batch sequencing were not performed.

### **Sequence processing and analysis**

Shotgun metagenomic reads were first trimmed and quality filtered to remove sequencing adapters and host contamination using Trimmomatic [24] and Bowtie2 [25], respectively, as part of the KneadData pipeline version 0.10.0 (<https://huttenhower.sph.harvard.edu/kneaddata/>). Recognizing that urine samples have lower biomass, we have taken steps to ensure results would not be affected by random contamination by rejecting samples below a threshold of a certain number of metagenomics reads (500,000, see below). This helps ensure removal of samples representing low abundance laboratory contaminants and robust sequence counts for downstream analysis to reduce batch-to-batch variation.

Relative abundances of different microbial taxa were calculated using Metaphlan3 [26]. Relative abundances of detected microbial metabolic pathways were determined using HUMAnN3 [27] and measured as Read Per Kilobase Million (RPKM) (number of aligned reads/kilobase gene size \* 1000000) to account for varying sizes of alignment targets. Relative abundances of virulence factors were measured by aligning KneadData processed reads against protein sequence databases of known bacterial virulence factors (<http://www.mgc.ac.cn/VFs/>) [28]. Sequences were aligned using the DIAMOND aligner [29] in blastx mode with default settings for sensitivity of alignment (> 60% identity). Abundance of virulence factor genes was also measured in RPKM.

### **Data analysis**

Analyses were performed in R and Python. Shannon diversity indices, PERMANOVAs on Bray-Curtis distances, and tests of homogeneity of betadispersion were calculated using the vegan package [30]. Non-metric MultiDimensional Scaling (NMDS) ordination plots were generated using the R phyloseq package [31].

Differential abundance of taxonomies across variables of interest were measured with Maaslin2 [32], using the compound-Poisson linear model method due to the zero-inflated nature of urinary microbiomes and default settings for statistical significance and the Benjamini-Hochberg correction for multiple comparisons. Basic cleaning of collected clinical data was performed in python JupyterNotebook with the pandas and numpy packages. Specifically, unstructured clinical data was filtered for relevant clinical data of interest (e.g. diagnoses and chart notes directly associated with UTI) and objective variables were converted to binary for machine learning analyses. Urinary symptoms were left out of machine learning models as these were used to distinguish between asymptomatic patients with positive UA (suggestive of ASB) and symptomatic patients (suggestive of UTI) in our study; thus, inclusion would have introduced obvious bias into the models. Graphs were generated using the R ggplot2 package [33]. Final figures were arranged using Adobe Illustrator and BioRender.

### **Machine learning based classifier analysis**

We examined seven approaches for handling high-dimensional tabular datasets that have previously been applied to microbiome data, as well as a neural network-based algorithm, which can show superior performance with large datasets [34]: Elastic Net Regression [35], Least Absolute Shrinkage and Selection Operator (LASSO)

regression [36], Ridge Regression [37], Support Vector Machine (SVM) regression [38], a Random Forrest Classifier [39] (RFC), TabNet [40], and XGBoost [41]. All approaches were implemented to predict symptomatic versus asymptomatic cases (“0” represents asymptomatic cases, and “1” represents symptomatic cases), utilizing features derived from clinical covariates which included the results of clinical testing, including the results of microscopic UA as well as pertinent past medical history listed in **Table 1**. This data was then combined with the results of microbiome analysis which included measurements of taxonomies, metabolic pathways, and known virulence factors represented in the sequencing data. This resulted in three datasets taxonomies + clinical covariates (761 features), metabolic pathways + clinical covariates (456 features), and virulence factors + clinical covariates (2171 features).

For in-depth explanations of the machine-learning based methods used in this analysis, please see our materials supplement. We examined model performances by F1 score, Area Under the Receiver-Operator Curve (AUC), Accuracy, Precision, Recall, and Specificity. The F1-score is a harmonic mean between precision and recall. This accounts for both prediction errors and the specific type of prediction error. We highlight the F1-score as our main model evaluation metric because the F1-score accounts for imbalanced data and extremes in either recall or precision while maximizing both. Thus, this metric is well suited for cases where both false positives and false

**Table 1.** Baseline study cohort demographics.

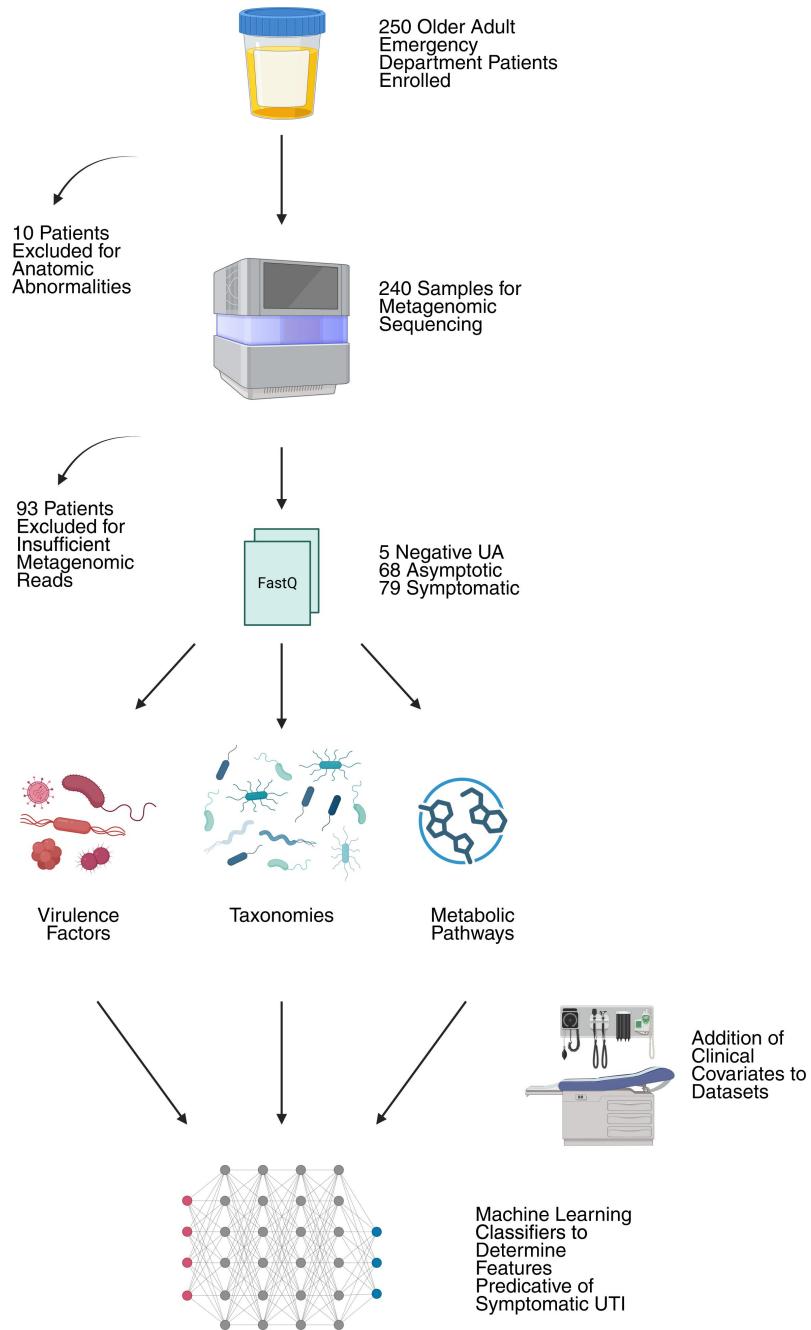
n	Asymptomatic Patients 68	Symptomatic Patients 79	p
Age (median [IQR])	80.00 [72.00, 87.25]	79.00 [73.00, 84.00]	0.408
Female (%)	56 (82.4)	61 (77.2)	0.572
White (%)	63 (92.6)	72 (91.1)	0.975
Black (%)	1 (1.5)	3 (3.8)	0.722
Other (%)	4 (5.9)	4 (5.1)	1
Ethnicity Hispanic (%)	4 (5.9)	2 (2.5)	0.545
Antibiotic Treatment in previous 90 days (%)	27 (39.7)	27 (34.2)	0.602
Treated in hospital with antibiotics for UTI (%)	57 (83.8)	66 (83.5)	1
Urinary Catheterization within 90 days (%)	5 (7.4)	11 (13.9)	0.313
Hospitalization within 90 days (%)	28 (41.2)	22 (27.8)	0.127
Resides in Private Home (%)	57 (83.8)	68 (86.1)	0.881
Resides in Assisted Living (%)	6 (8.8)	4 (5.1)	0.566
Resides in Nursing Home (%)	5 (7.4)	7 (8.9)	0.975
Surgery within 90 Days (%)	10 (14.7)	14 (17.7)	0.788
No Diagnosis of Dementia (%)	54 (79.4)	69 (87.3)	0.283
History of Alzheimer's Dementia (%)	1 (1.5)	3 (3.8)	0.722
History of Vascular Dementia (%)	2 (2.9)	0 (0.0)	0.412
History of Dementia Not Otherwise Specified (%)	9 (13.2)	6 (7.6)	0.394
Baseline Incontinence of Urine (%)	32 (47.1)	40 (50.6)	0.79
History of Diabetes (%)	30 (44.1)	27 (34.2)	0.288
Active Cancer Diagnosis (%)	15 (22.1)	29 (36.7)	0.08
On Immunosuppressive Medication (%)	6 (8.8)	8 (10.1)	1
History of COPD (%)	7 (10.3)	23 (29.1)	0.009*
Baseline Incontinence of Stool (%)	7 (10.3)	12 (15.2)	0.525
Admitted to Hospital (%)	44 (64.7)	49 (62.0)	0.869
Placed in Observation (%)	22 (32.4)	23 (29.1)	0.806
Discharged from ED (%)	2 (2.9)	7 (8.9)	0.251

negatives are undesirable. AUC provides a balanced measure between the true positive rate and false positive rate, describing how well a model distinguishes between positive and negative classes. Recall is the rate of true positive prediction. Precision is how often a true positive result is predicted. Specificity is how often a true negative result is predicted. Accuracy measures the proportion of both true positives and true negatives among the total number of cases, assessing the model's overall correctness. F1 score identifies positive cases while equally penalizing false positives (FP)

and false negatives (FN), thereby balancing concern for both precision and recall. A summary of these metrics for each of the different models is shown in Table S2.

## Results

The overall study schema is shown in [Figure 1](#). Of 250 enrolled patients, 10 were excluded due to anatomic abnormality (suprapubic catheter, urostomy, etc.). An additional 93 were excluded for insufficient metagenomic reads. Although this represents a large portion



**Figure 1.** Study overview and analysis scheme.

of the cohort, this exclusion ensured that low abundance contaminants that could be introduced by handling and batch effects from different sequencing runs did not overly impact downstream analysis. Of the remaining cohort, 147 returned positive UAs with or without urinary symptoms while 5 had negative UAs, all without symptoms.

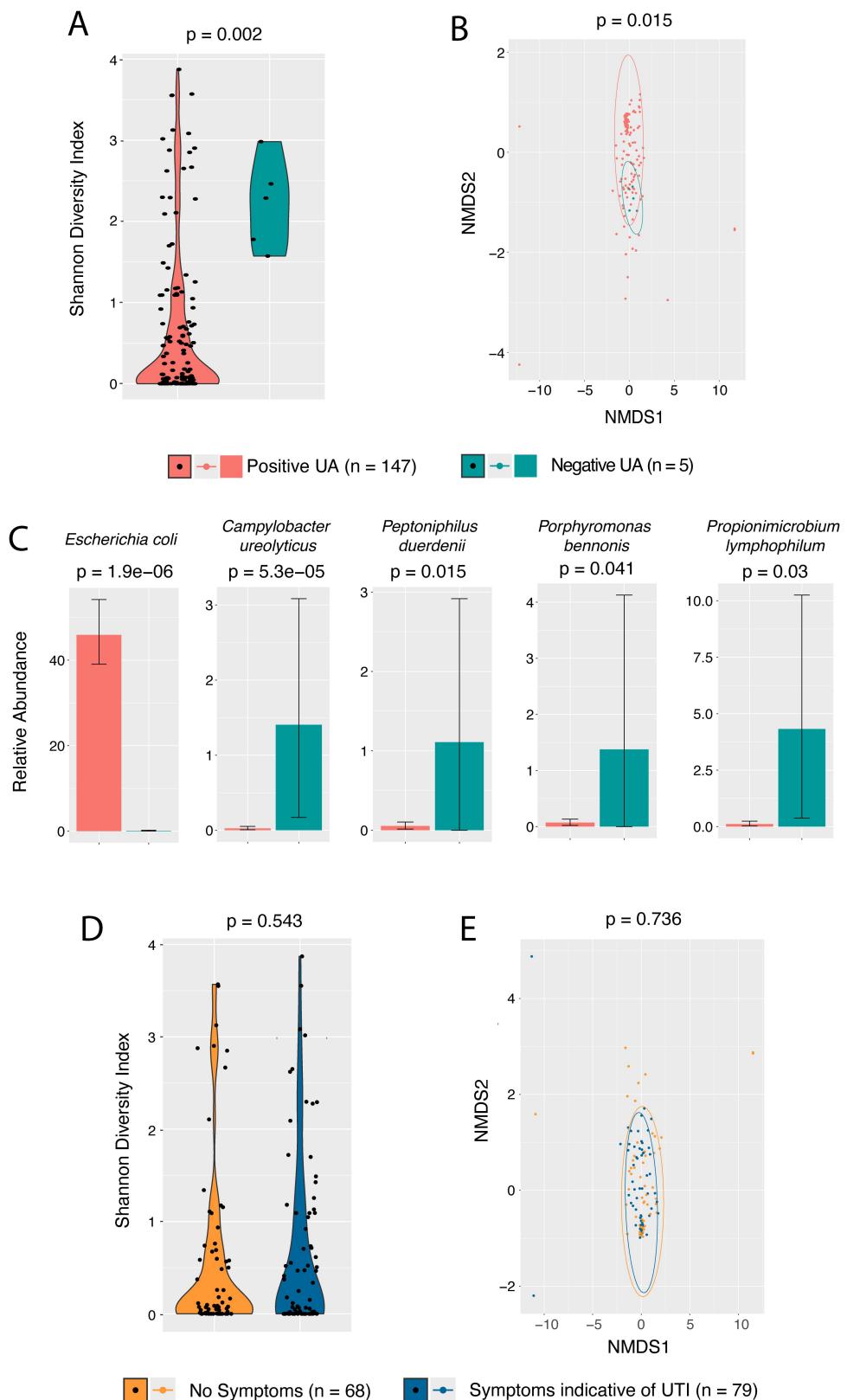
When examining negative UA vs positive UA microbiomes, negative UA communities showed higher richness (number of species in each microbiome) (48.4 mean species negative UA versus 19.7 positive UA,  $p = 0.01$  for ANOVA on richness), and evenness (similar abundance across all species detected) (mean Pielou's evenness 0.59 negative UA and 0.3 positive UA,  $p = 0.018$  for ANOVA on evenness). Alpha diversity, a combined metric reflecting richness and evenness within individual samples, as measured by Shannon index (Figure 2(A)), was significantly higher in negative UA microbiomes (median Shannon index 2.286 for negative UA versus 0.08633 in positive UA samples,  $p = 0.002$  Mann-Whitney U-test). Examining beta diversity, a metric of how much microbiome compositions differ between groups, we found significant centroid distances in Non-metric MultiDimensional Scaling (NMDS) ordination plots (Figure 2(B)) between positive UA and negative UA groups (PERMANOVA on Bray-Curtis distance  $p = 0.0033$ ), variances were not significantly different (ANOVA on beta-dispersion  $p = 0.778$ ). Utilizing the Microbiome Multivariable Association with Linear Models (Maaslin2) package [32], we found that that positive UA samples contained more *Escherichia* ( $p = 1.9 \times 10^{-6}$ ). Other organisms detected in much smaller amounts included *Campylobacter ureolyticus* ( $p = 5.3 \times 10^{-5}$ ), *Peptoniphilus duerdenii* ( $p = 0.015$ ), and *Propionibacterium lymphophilum* ( $p = 0.04$ ) were higher in the negative UA samples (Figure 2(C)).

We next examined positive UA samples as symptomatic versus asymptomatic cases on the basis of signs and symptoms of UTI (68 categorized as asymptomatic, 79 as symptomatic). The baseline characteristics of symptomatic and asymptomatic patients were very well balanced with respect to demographics, recent antibiotic treatment, hospitalizations, primary residence, baseline incontinence of stool or urine, and dementia diagnoses (Table 1). The only medical history that varied between groups was a history of Chronic Obstructive Pulmonary Disease (COPD) ( $p = 0.009$ , chi-square). The association between UTI and COPD among older adults has been reported in the past [42]. Individuals with COPD are thought to be at increased risk of UTI due to treatment with anti-cholinergic medications used to manage COPD causing urinary retention [43]. We also examined if any features of

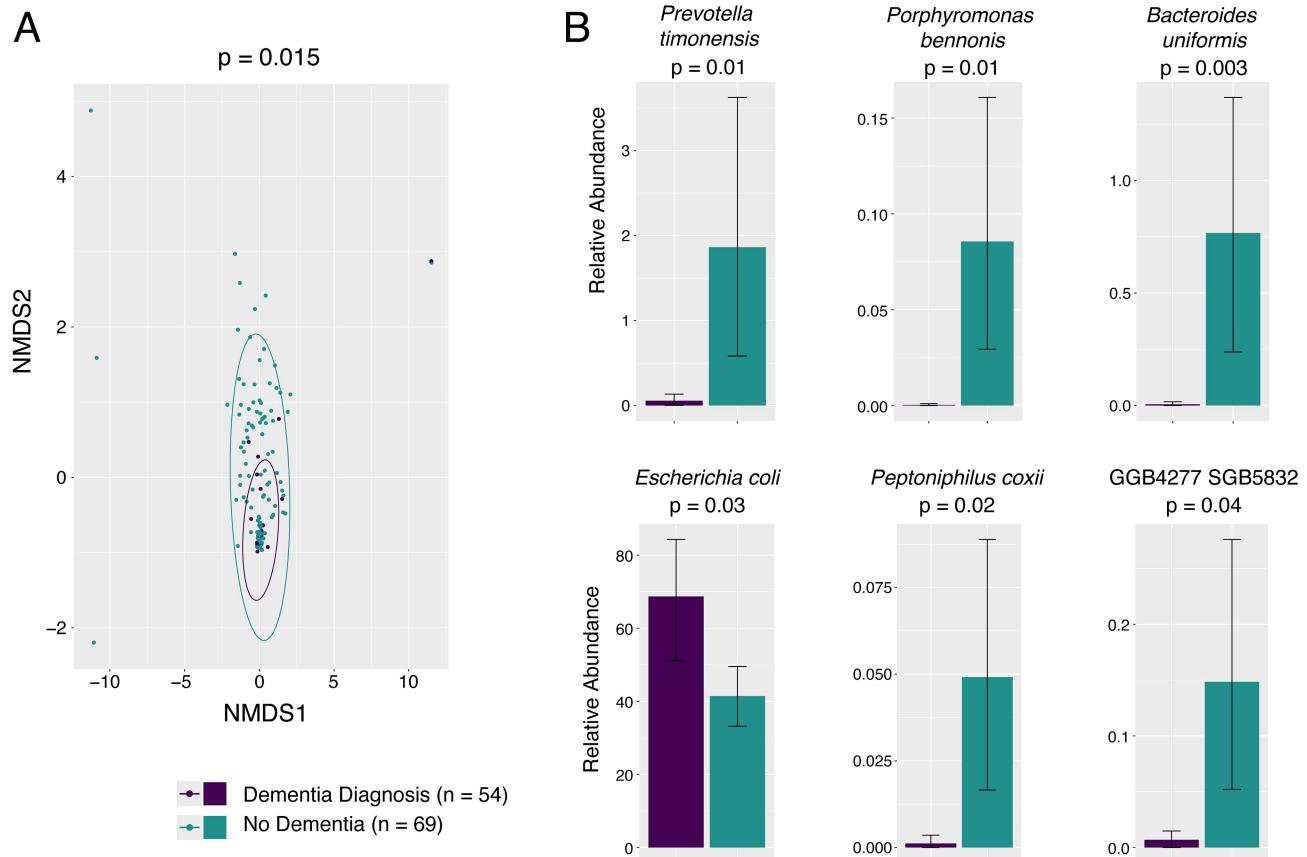
the microscopic UA typically used to diagnose infection (white blood cells/hpf, red blood cells/hpf, presence of nitrites, qualitative presence of bacteria) varied across symptomatic and asymptomatic cases. There were no significant differences in any of these metrics when examined across symptomatic vs asymptomatic cases (Supplemental Figure 1).

We next examined broad metrics of microbiome composition across these groups. Alpha diversity did not differ significantly between asymptomatic and symptomatic cases (median Shannon index in 0.0819 versus 0.116,  $p = 0.543$  Mann-Whitney U-test) (Figure 2(D)). No significant difference in the centroid distances between asymptomatic and symptomatic cases was seen (PERMANOVA on Bray-Curtis distance  $p = 0.736$ ) (Figure 2(E)). We also explored Bray-Curtis distance between microbiomes associated with clinical covariates of interest, including urinary catheterization within the prior three months, sex, living in the community versus residing within a nursing home or assisted living apartment, hospitalization or surgery within the last three months, presence of dementia, incontinence of stool or urine, and antibiotic treatment for any indication within the last three months, the complete results of which are listed in Table S1. Beta-diversity did not differ when examining microbiomes from male and female participants, which was also true in our previous study of the urinary microbiome of nursing home residents [44], and may reflect age-related changes in the microbiome that are shared among men and women. When comparing participants with and without a diagnosis of dementia, there were significant differences in beta diversity (PERMANOVA on Bray-Curtis  $p = 0.015$ ,  $n = 123$  without dementia,  $n = 24$  with some dementia diagnosis), although the beta dispersion did vary significantly between these groups as well (ANOVA on beta-dispersion  $p = 0.0016$ ) which could be somewhat responsible for this finding (Table S1). The NMDS ordination plot of Bray-Curtis distances between these two groups is shown in Figure 3(A). We also explored which taxa were associated with dementia diagnosis using Maaslin2. Six species associated with a diagnosis of dementia. *E. coli* was returned as having a positive association with the diagnosis of dementia ( $p = 0.003$ ) while five other taxa present at lower relative abundance showed a negative association with dementia. Mean relative abundances of these taxa across the two groups and significance of the association are shown in Figure 3(B).

We did explore differences in clinical covariates among study participants with and without a diagnosis of dementia, such as age, recent urinary



**Figure 2.** Community level microbiome metrics comparing positive UA versus negative UA, and asymptomatic vs symptomatic cases. (A) Subjects with a negative UA had significantly higher alpha diversity of their urinary microbiome (median Shannon index 2.286 for control versus 0.08633 for positive UA,  $p = 0.002$  Mann-Whitney U-test). (B) NMDS ordination plot of positive UA and Negative UA microbiomes shows that the positive UA cohort had substantially more variation between samples compared to negative UA. (C) Differential abundance of taxa that vary significantly between negative UA and positive UA using Maaslin2. Shown is the mean and standard deviation of relative abundance of each taxa as well as p-value associated with coefficient in the model. *E. coli* represented



**Figure 3.** Urinary microbiome among subjects with dementia compared to those without dementia. (A) NMDS ordination plot of Bray-Curtis distance between urinary microbiomes among subjects with and without a diagnosis of dementia (PERMANOVA on centroid distances  $p = 0.015$ ). (B) Mean relative abundance of taxa in urinary microbiomes found to be predictors of dementia diagnosis by Maaslin 2. Shown is mean and standard deviation of relative abundance and p-value of coefficients determined by Maaslin2. *E. coli* shows substantial higher relative abundance in individuals with dementia, other taxa were present at lower relative abundance and higher in individuals without dementia.

catheterization, hospitalization and prior treatment with antibiotics directed against UTI. These results are shown in Table 2. Patients with a diagnosis of dementia were older ( $p = 0.047$ , Mann-Whitney U test), but no other clinical covariate we hypothesized could contribute to urinary microbiome dysbiosis varied significantly differently across the groups, although there was a trend towards additional antibiotic treatment for UTI in the past (Chi-squared  $p = 0.096$ ).

An advantage of metagenomic sequencing is being able to analyze microbiome features beyond taxonomy. Thus, we generated datasets of relative abundance of genes within known metabolic pathways [27] and relative abundance of known virulence factors by counting

reads aligned against the VFDB database of protein sequences associated with known virulence factors [28]. We trained machine-learning based classifier models on three multimodal datasets: available clinical data (without specific signs and symptoms of UTI used to classify cases) combined with either relative abundance of taxa, RPKM of metabolic pathways, or RPKM of known virulence factors. The neural network-based TabNet showed the most consistent results across all performance metrics and datasets. Some penalized regression models such as LASSO and ridge regression showed higher mean F1 scores, but at the cost of specificity (Table S2). Most of the models we tested return ranked importance of which features of the

---

a substantial portion of mapped microbial reads from those with positive UA compared to negative, but the relative abundance of some organisms was higher in the negative UA samples. (D) Shannon alpha diversity index did not vary across asymptomatic vs symptomatic cases (median Shannon index in 0.0819 versus 0.116,  $p = 0.543$  Mann-Whitney U-test). (E) NMDS ordination plot of Bray-Curtis distance between asymptomatic vs symptomatic cases, significant overlap between the two states is seen.

**Table 2.** Examination of clinical factors across patients with and without dementia.

n	Any Dementia Diagnosis 24	No Dementia 123	p
Age (median [IQR])	84.00 [77.00, 87.50]	79.00 [72.00, 86.00]	0.047*
Female (%)	20 (83.3)	97 (78.9)	0.826
White (%)	21 (87.5)	114 (92.7)	0.659
Black (%)	1 (4.2)	3 (2.4)	1
Other (%)	2 (8.3)	6 (4.9)	0.849
Hispanic (%)	1 (4.2)	5 (4.1)	1
Antibiotics within 90 Days (%)	11 (45.8)	43 (35.0)	0.436
Antibiotics for UTI specifically within 90 Days (%)	8 (33.3)	20 (16.3)	0.096
Urinary Catheterization within 90 Days (%)	4 (16.7)	12 (9.8)	0.525
Hospitalization within 90 Days (%)	9 (37.5)	41 (33.3)	0.874
Surgery within 90 Days (%)	4 (16.7)	20 (16.3)	1
Baseline Incontinence of Urine (%)	11 (45.8)	61 (49.6)	0.909
Baseline Incontinence of Stool (%)	3 (12.5)	16 (13.0)	1
Community Dwelling (%)	19 (79.2)	106 (86.2)	0.57
History of Chronic Kidney Disease (%)	10 (41.7)	40 (32.5)	0.529

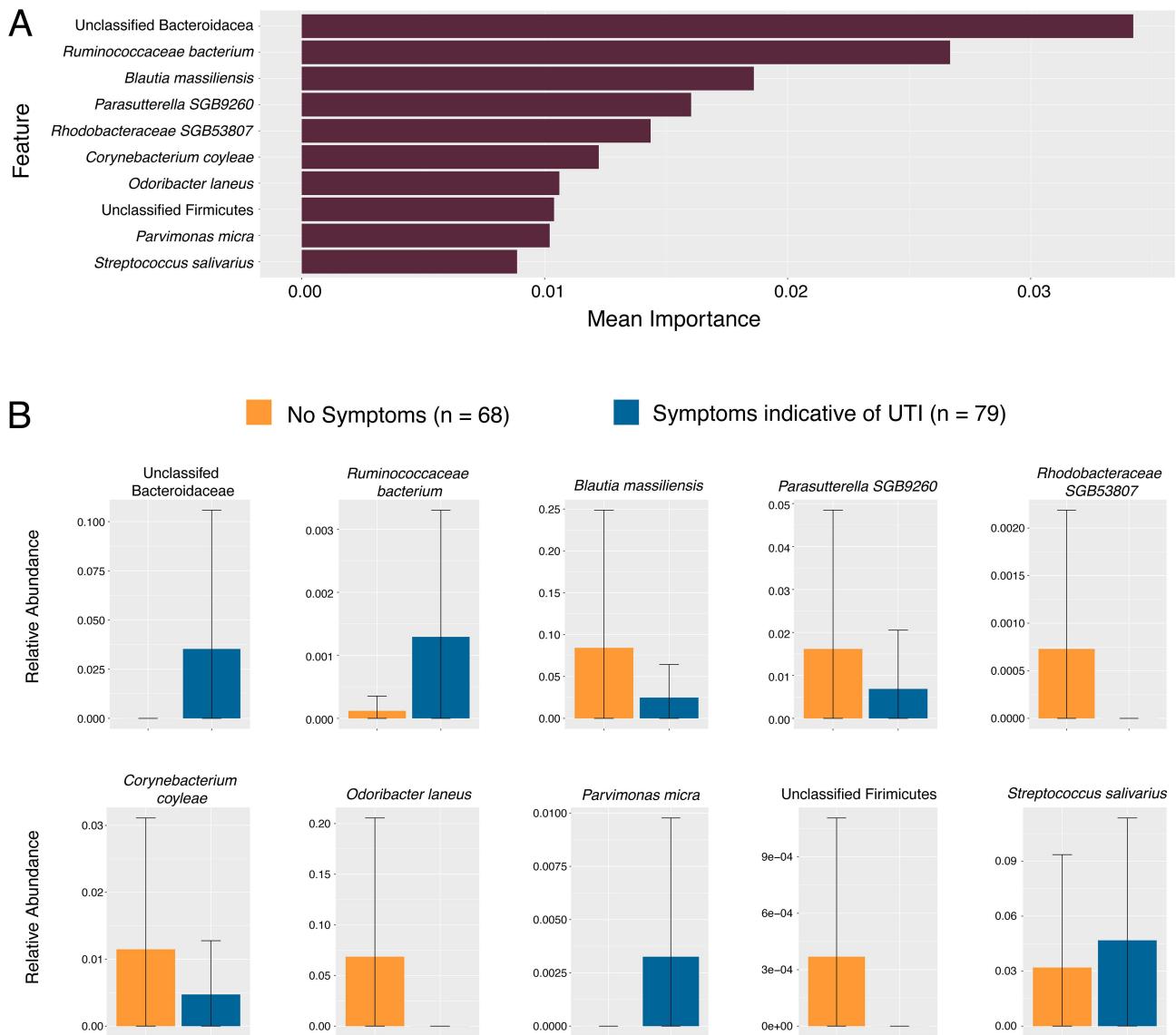
modeled datasets contributed most to the classification outputs (symptomatic versus asymptomatic). In most models, microbiome features (taxa abundance, metabolic pathways, or virulence factors) were ranked as more important than non-symptom clinical inputs. We selected TabNet as our final model due to its consistent performance across all datasets in this study. Compared to other classifiers, TabNet achieved the most balanced trade-off between sensitivity (recall) and specificity. Figures 4–6 (top panels) show the ranked importance from the TabNet models.

Examining the important features in the TabNet model built on relative abundance of different taxa, the relative abundance of a species of an unclassified *Bacteroidaceae* (mean relative abundance 0% in asymptomatic versus 0.035% in symptomatic), was the most important feature, followed by the abundance of *Ruminococcaceae bacterium*, with higher relative abundance in our symptomatic group (mean relative abundance 0.00012% asymptomatic versus 0.0013% in symptomatic). The abundance of *Blautia massilensis* (mean relative abundance 0.085% versus 0.025%) and a *Parasutterella* (mean relative abundance 0.016% versus 0.00012%) taxa were the next two most important features and showed higher abundance in urinary microbiomes from individuals classified as asymptomatic (Figure 4). In the model built on metabolic pathway abundance, summarized in Figure 5, the two most important features were the abundance of the Pentose-phosphate pathway (mean RPKM 2730 in symptomatic versus 2567 in asymptomatic), and a clinical covariate, the level of serum creatinine, both higher in symptomatic patients (Figure 5). The pathways for homolactate fermentation, and menaquinone biosynthesis were also higher in microbiomes of symptomatic patients (mean RPKM 1647.5 versus 1131.5 and mean RPKM 1210.5 versus 1131.5 respectively). Pathways for fatty acid beta-oxidation, nicotinamide (NAD) salvage, and

glycogen degradation were more abundant in asymptomatic cases (mean RPKM 25.3 versus 12.3, 103 vs 80, and 812.5 vs 687.5 respectively). Another clinical covariate, a positive urine culture for *Klebsiella* was slightly higher in asymptomatic cases. In the model built on relative abundance of virulence factors, summarized in Figure 6, the most important features were a type II secretion protein (genbank ID WP\_016352169) (mean RPKM 58.1 in symptomatic versus 71.8 in asymptomatic), a component of flagellum assembly (genbank ID WP\_015444575) (mean RPKM 38.9 versus 40.3) and a type IV secretion system effector protein (Genbank ID WP\_010948683) (mean RPKM 1.5 vs 0.83).

## Discussion

There appears to be a clear distinction in our cohort between the urinary microbiome from patients with a positive microscopic UA and those whose UA do not show any findings consistent with infection (negative UA). These include much higher alpha diversity, higher richness, and evenness for negative UA and clear differences in beta diversity between negative and positive UA. This finding is consistent with studies of the microbiome in the gastrointestinal tract [45], which show that in general, a more diverse microbiome is reflective of a healthy microbiome; in this way a UA with findings consistent with infection and lower in microbial compositional diversity could be considered a dysbiotic or unhealthy state. The presence of diverse commensal organisms could potentially be harnessed as testing for a “healthy” urinary microbiome and could be targeted in future studies to determine if they perform some active function in preventing pathogen colonization. Alternatively, the presence of inflammation within the urinary tract could also drive dysbiosis. Previous work has shown that in both ASB and UTI, there are higher abundances of proteins associated with



**Figure 4.** Top Ten features of TabNet model on relative abundance of taxonomies and clinical covariates. (A) Graphed mean importance of each feature in the model. (B) Mean relative abundance of top 10 most important taxa in predicting asymptomatic versus symptomatic cases in the TabNet model, error bars represent standard deviation of the mean. No formal statistical tests are performed on these features aside from TabNet analysis and as such no p-values are reported.

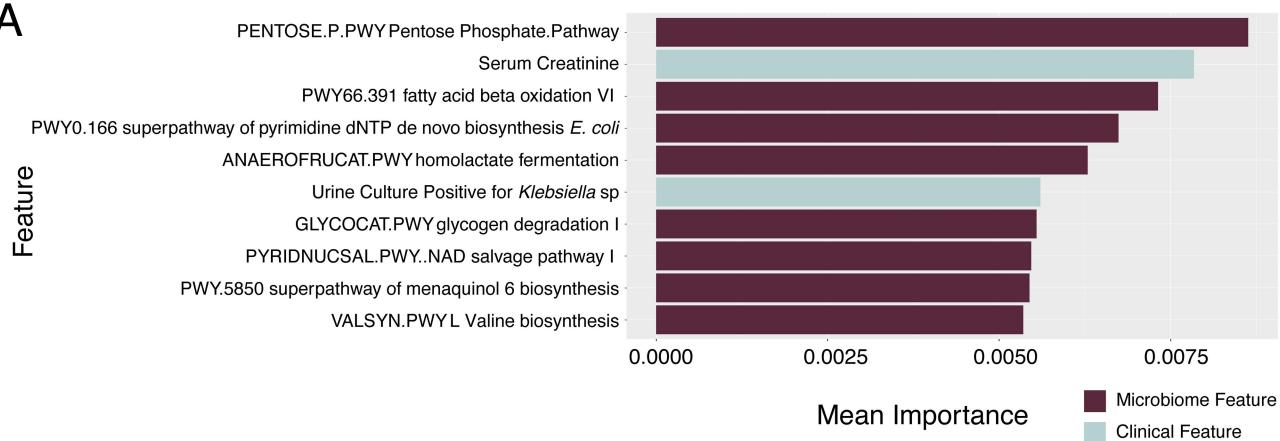
neutrophil activation, degranulation, and chemotaxis in the urine of individuals with evidence of infection [46].

Of particular interest is the overwhelming presence of *E. coli* in those with positive UA, suggesting this organism is much more common in subjects with a positive UA, whether they have symptoms or not. *C. ureolyticus*, *P. durendeni*, and *P. lymphophilum* were more abundant in negative UA microbiomes. These organisms have been variously reported to be potential human pathogens [47–49], but their presence here in the urine of patients with negative UAs indicates they may be beneficial commensal organisms. Alternatively, they may become pathogenic in the setting of dysbiosis or specific perturbation of commensal community structure. Thus, normal colonization

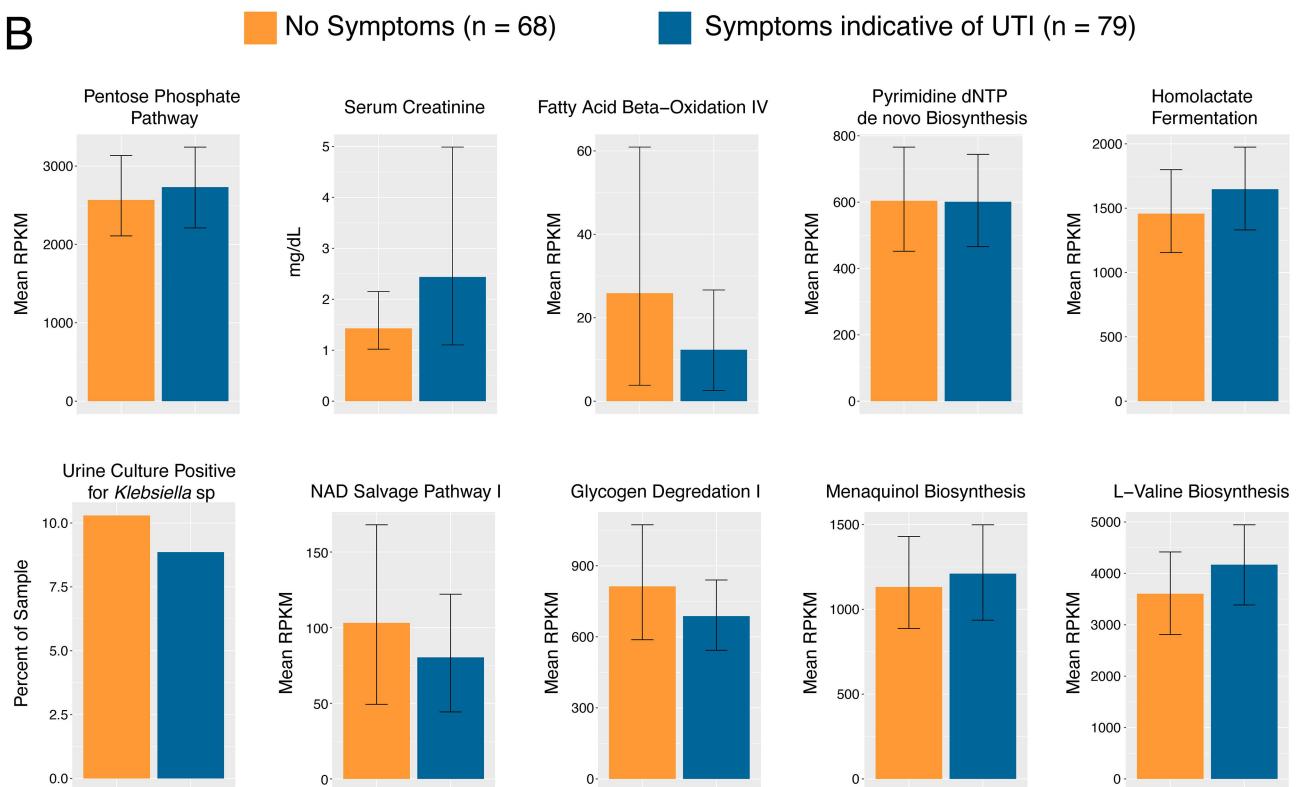
of specific bacteria within a healthy urinary microbiome may exist along a continuum wherein dysbiosis of specific species may precede ASB and UTI and symptoms present in context with immune activation and an altered uromicrobiome.

An unexpected result was a difference in beta-diversity of the urinary microbiome between patients with dementia versus those without. Although beta-dispersion did vary between these groups, PERMANOVA is considered to be a test that is least likely to be affected by heterogeneity in beta-dispersion between groups [50]. When examining what features were associated with a dementia diagnosis using Maaslin2, the potential pathogen, *E. coli*, was detected more frequently and in higher

A



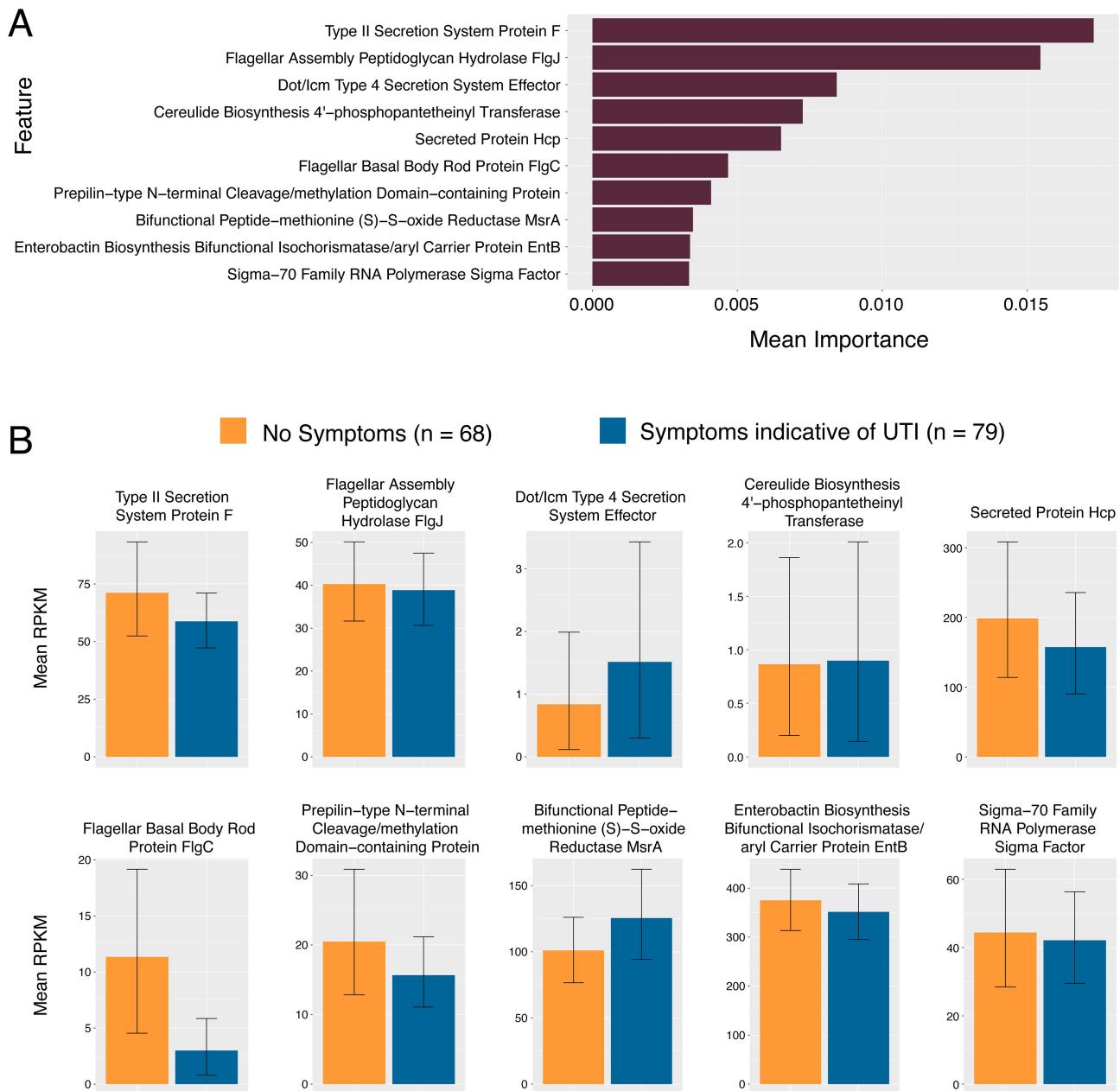
B



**Figure 5.** Top Ten features of TabNet model on RPKM of bacterial metabolic pathways and clinical covariates. (A) Graphed mean importance of each feature in the model. (B) Graphed mean RPKM and standard deviation of reads mapped to established metabolic pathways split by asymptomatic versus symptomatic cases. For serum creatinine, this is measured mg/dL abstracted from clinical data. For positive urine culture for *Klebsiella* species, graphed is the percent positive samples from asymptomatic vs symptomatic cases respectively. No formal statistical tests are performed on these features aside from TabNet analysis and as such no p-values are reported.

abundance in those with a dementia diagnosis, while higher abundances of *Prevotella timonensis*, *Porphyromonas bennonis*, *Bacteroides uniformis*, *Peptoniphilus coxii*, and an uncharacterized taxa were negatively associated with dementia. While some of these organisms may be opportunistic pathogens [51–53], their presence at relatively low abundance suggests they are likely commensals in this context, and

overgrowth of *E. coli* may represent dysbiosis among individuals with dementia. We did explore non-symptom clinical factors among patients with dementia that could potentially explain the observed dysbiosis such as recent antibiotic treatment, hospitalization, urinary catheterization, or residing in the community versus a nursing home. Although patients with dementia were older (median age 84 with dementia, 79 without,  $p = 0.047$



**Figure 6.** Top Ten features of TabNet model on RPKM of known virulence factor genes and clinical covariates. (A) Graphed mean importance of each feature in the model. (B) Mean RPKM and standard deviation of reads that mapped to known virulence factors split between asymptomatic versus symptomatic cases. No formal statistical tests are performed on these features aside from TabNet analysis and as such no p-values are reported.

Mann-Whitney U test), no other clinical covariate was significantly associated with dementia (Table 2). How dementia may be driving this dysbiosis or vice versa could be investigated in the future.

On initial examination, the most striking feature of our cohort of subjects with positive UAs is the very high number of asymptomatic cases (68 cases or 46%). These patients should not have had urinalysis obtained at all if clinical guidelines were strictly adhered to. The microbial community of individuals with a positive UA

was remarkably similar whether they had signs and symptoms of UTI or not. No diversity metrics showed significant differences. We have used machine learning-based classifiers to explore multimodal microbiome data to find meaningful microbiome features associated with clinical conditions or outcomes in the past and applied it to this dataset [54,55]. We employed a wide array of popular machine learning-based classifier algorithms, of these, the neural-network based TabNet model performed most consistently across metrics.

TabNet's architecture supports instance-specific feature selection through its sequential attention mechanism. Unlike Lasso, Ridge, or SVM, which use a fixed set of features across all samples, TabNet adjusts the relevance of features for each case. This flexibility is particularly important in microbiome-based classification, where patient-specific microbial signatures often vary substantially. As shown in Table S2, TabNet has the most consistent performance metrics with superior specificity, consistent recall, sample-level feature selection, and low variance across seeds.

Although we hypothesized that microbiome features would reliably distinguish symptomatic versus asymptomatic patients and that these features could subsequently be examined for use as enhanced diagnostics for true UTI, we acknowledge that the performance of these models is not reliable enough to be used in clinical testing to discriminate between potential ASB and UTI. The high number of asymptomatic cases and the modest performance of our in-depth microbiome analysis highlight that antibiotic stewardship campaigns that focus on education and adherence to guidelines remain the most likely interventions to reduce inappropriate testing and antibiotic treatment. However, the high importance of microbiome features in our models provides biological insight into the importance of the urinary microbiome in these two states, and we believe some broad themes emerge.

In our model built on taxonomies, we observed that the most important feature was the relative abundance of an uncharacterized *Bacteroidaceae* strain, which was more abundant in microbiomes classified as symptomatic. This family is typically anaerobic and associated with the gut microbiome. Its presence here may indicate ascending colonization of microbiota from the gut which is suggested to be the source of some uropathogens [56]. Another taxa that showed higher abundance in symptomatic microbiomes was a strain of *Ruminococcaceae*, which is a potential pathobiont that has been associated with the urinary microbiome of individuals with bladder cancer [57] and was positively correlated with lower urinary tract symptoms in men [58]. Another taxa with higher abundance in microbiomes classified as asymptomatic include a *Parasutterella* species, a taxa that is more typically observed in the gut microbiome [59], but interestingly, is depleted in individuals who suffer from recurrent UTIs [56], suggesting it maybe a protective commensal. *B. massilensis* was also more abundant among asymptomatic patients, members of this family and genus have been negatively associated with urinary symptoms in men [58], suggesting it may be another protective commensal.

When examining the model built on relative abundance of microbial metabolic pathways and non-symptom clinical features, important features may be suggestive of a role for specific patho-adaptations in eliciting symptoms. Features more abundant in microbiomes from symptomatic patients included the pentose phosphate pathway, homolactate fermentation, and menaquinol biosynthesis. Considering these in the context of infection, the pentose phosphate pathway in *E. coli* is involved in generating reducing potential and building blocks for biomolecules [60] and is upregulated in *E. coli* in the setting of UTI [61]. Oxygen tension in urine can be quite low and it is likely that anaerobic fermentation (such as homolactate) pathways are important virulence factors for UPEC [62]. Menaquinone is specifically involved in anaerobic metabolism and UPEC mutants in this pathway show decreased fitness in models of urinary tract infection [63]. Important microbial pathways with higher abundance in microbiomes classified as asymptomatic included fatty acid beta oxidation, NAD salvage pathway, and glycogen degradation. Fatty acid biosynthesis, as opposed to degradation, has been observed to be upregulated in *E. coli* isolated from human UTI specimens [64]. *E. coli* seems to rely on nucleotide salvage for planktonic growth in urine, but requires *de novo* synthesis during invasive infection [60]. Glycogen metabolism was seen to be important in a mouse model of *E. coli* intestinal colonization [65], but utilization of other carbon sources seems to be more important in causing UTI [66]. Overall, pathways associated with colonization show higher abundance in microbiomes classified as asymptomatic, whereas those that seem to be important in rapid growth are more abundant in symptomatic UTI microbiomes. Indeed, the ability to grow rapidly in the relatively anaerobic host environment may be the decisive factor in causing infection [67] as opposed to simply colonizing the urinary tract. Of note, two clinical variables appeared among the top ten features in this model, including a higher serum creatinine in symptomatic individuals, which may reflect decreased kidney function in individuals suffering from UTI, and the percentage of patients with positive urine cultures for *Klebsiella*, suggesting that this organism may more commonly be associated with asymptomatic colonization.

When examining the model built on relative abundance of virulence factors and clinical features, two important themes were secretion systems and bacterial flagellum components. Genes associated with flagellum assembly (*flgJ* and *flgC*) were more abundant in asymptomatic microbiomes. The flagellum is a structure involved in bacterial motility, and in a mouse model

of cystitis, this was seen to be important in persistence in the bladder [68]. Type II and Type IV secretion systems are ubiquitous among Gram negative pathogens and may have a role in the pathogenesis of UPEC [69]. Interestingly, the type II secretion system component *gspF* was more abundant in asymptomatic microbiomes where as a Type IV secretion effector was more abundant in symptomatic microbiomes (mean RPKM 1.5) versus asymptomatic (mean RPKM 0.84). These systems are pathogenicity factors for a wide range of Gram-negative pathogens [70], but their role in UTI has not been established.

It is important to note that the classifiers used here return feature importance in predicting outcomes as defined by our clinical and laboratory criteria. However, they do not inherently return the direction of the association (e.g. whether a higher or lower abundance, presence or absence of a feature is predictive) or a traditional statistical confidence or p-value associated with feature importance. Neural networks can consider features in combination; so the abundance in each individual sample may only be predictive as part of a combination of other feature patterns considered in the model. In our discussion, we have simplified the analysis of predictive features by examining their abundance in one state versus another to attempt to make inferences regarding biologic processes which may underly the development of symptomatic UTI. These features should be considered together as part of a pattern distinguishing microbiomes from asymptomatic and symptomatic patients and we have tried to highlight unifying themes within results.

### **Strengths and weaknesses**

This was a well-balanced cohort that incorporated clinical covariates and microbiome features into our analysis and showed significant differences between balanced groups (symptomatic versus asymptomatic, dementia versus no dementia). Strengths also included microbiome measurements taken from point-of-care samples, inclusion of laboratory testing, and detailed surveys of current signs and symptoms for as accurate of a distinction between UTI and ASB diagnosis as possible on the basis of all available data in the acute care setting. We also were able to include patients who lacked capacity to consent, a group for which clinicians face difficulty in obtaining history of UTI symptoms and represent a greater diagnostic challenge. Although our final F1 scores from machine learning models were modest, all models trended towards positive ( $> 0.5$ ) predictions.

A weakness of this study includes its relatively small final cohort (68 asymptomatic patients, 79 with symptoms, and 5 patients with negative UA). As such, results of our analysis of negative UAs may not be widely generalizable as this is based on the small sample size of five individuals. Many patients/samples had to be discarded because of very low metagenomic content, which was necessary to ensure robust microbiome data for analysis. In microbiome analysis of low biomass sites like urine and airway, it is possible that contamination by handling or introduced by lab equipment can bias results [71]. Although we did not perform negative control sequencing samples, we did not detect as significant any taxa commonly associated with skin or environmental microbiome. Also, issues with contamination can be avoided with higher biomass inputs, which we effectively instituted by using a minimum metagenomic sequencing cut off [72]. Analyzing samples with minimum metagenomic content also helps prevent bias from different sequencing batches [73].

Another limitation was the use of convenience samples, which had to be evaluated and ready for research staff to obtain during regular business hours. It is possible that some unmeasured bias for certain patient types could be introduced by this practice and the results may not be generalizable across patients presenting at any time. Nonetheless, our measured demographic factors were well balanced across asymptomatic and symptomatic patients. Another weakness and potential explanation for our modest F1 scores may be that the patients who may have only had nonspecific symptoms, such as fever and altered mental status, may have had an alternative final diagnosis that was not captured by our 14-day chart review. It is also possible that individuals with dementia would be considered asymptomatic more frequently as they may have limited ability to report symptoms such as dysuria. However, our cohorts were balanced with respect to dementia diagnosis, so we do not believe that there was a bias towards classifying these patients as asymptomatic. Non-specific signs and symptoms like fever or altered mental status in the setting of a positive UA would usually be considered diagnostic and 84% of such cases in our cohort were indeed treated with antibiotics directed against UTI specifically. The IDSA guidelines do recommend that additional testing be performed for alternative explanations in such cases before initiating treatment for UTI [13], but our data does not capture what, if any, the results of those additional tests might be. If an alternative explanation was responsible for a patient's symptoms, some cases could be later classified more accurately as asymptomatic for machine learning model training. Despite these challenges, we were able to describe significant

community level differences between microbiomes present in positive versus negative UAs, as well as unexpected community differences in patients with and without dementia. Our machine-learning based analysis did identify microbiome features primarily as predictive of asymptomatic versus symptomatic cases, suggesting that there are features within uromicrobiome that are associated with symptoms in UTI.

## Conclusions

This metagenomic analysis of older adult ED patients with suspected UTI highlights the continued difficulties of diagnosing and treating older adult patients for UTI, especially in the acute care setting. Nearly half of our participants with a positive UA did not have signs and symptoms that would qualify for testing for UTI based on IDSA and Massachusetts Department of Public Health guidelines (and thus should not have been tested in the first place). While this can be regarded as a failure of strict guideline adherence on the one hand, on the other hand, the results of these asymptomatic UAs may give us insight about a previously underappreciated biologic process: the transition from a healthy to a dysbiotic uromicrobiome and further, the transition from ASB to true clinical infection. Indeed, clear definitions between these transition points and states are still undefined. Machine learning-based methods are becoming more prevalent in the analyses of complex microbiome data, and we have attempted to apply it here to divulge insight into what distinguishes ASB from UTI from a clinical perspective based on acute care setting information. Our analyses suggest some functional uromicrobiome characteristics (genes favoring colonization activities versus virulence activities) may aid in distinguishing these two states and may offer clues into the biologic mechanisms that result in physiologic symptoms of UTI.

## Acknowledgements

The authors thank the administrators Clinical Renal Laboratory at UMass Memorial Medical Center, they're willing to work with us allowed us to obtain the same urine samples for metagenomic analysis as were provided for clinical testing and allowed us to carry out study with minimal impact on any patient care teams.

ESB was responsible for study design, oversight of study conduct and data analysis. CS, LC, AL, OA, were responsible for subject consent and data collection. LP and TF were responsible for urine sample acquisition. DVW was responsible for sample preparation for sequencing and initial data processing. AZ, ZH, and VB were responsible for machine-learning based analysis. JPH and BAM served as mentors for this work providing guidance on study design and execution, data analysis, and manuscript preparation. All authors have

read and given approval for the written manuscript presented here.

## Author contributions

CRediT: **Evan S. Bradley:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; **Celina Stansky:** Investigation, Methodology, Project administration, Writing – original draft; **Abigail L. Zeamer:** Data curation, Formal analysis, Software, Writing – original draft; **Ziyuan Huang:** Formal analysis, Software, Visualization; **Lindsey Cincotta:** Investigation, Project administration; **Abigail Lopes:** Investigation, Project administration; **Linda Potter:** Investigation, Methodology; **Theresa Fontes:** Investigation, Methodology; **Doyle V. Ward:** Data curation, Formal analysis, Methodology, Resources, Software; **Vanni Bucci:** Methodology, Resources, Software, Supervision, Validation, Writing – original draft; **Beth A. McCormick:** Resources, Supervision, Writing – original draft; **John P. Haran:** Funding acquisition, Methodology, Project administration, Resources, Supervision.

## Funding

This work was supported by the Early Career Development Award from the Emergency Medicine Foundation (EMF) application number [EMF20\_ECRD\_00027].

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Data availability statement

Sequence files and metadata are available in the NCBI's Short Read Archive (SRA) under the project number PRJNA120898 by following the following link:  
[https://www.ncbi.nlm.nih.gov/bioproject?LinkName=sra\\_bio\\_project&from\\_uid=36843498](https://www.ncbi.nlm.nih.gov/bioproject?LinkName=sra_bio_project&from_uid=36843498)

## Protection for human subjects

The protocols used for this study adhered to the principles outlined in the Declaration of Helsinki and were reviewed and approved by the UMass Chan Institutional Review Board, Docket # H00014312.

## ORCID

Evan S. Bradley  <http://orcid.org/0000-0001-7704-6371>  
 Abigail L. Zeamer  <http://orcid.org/0000-0002-3600-4519>  
 Ziyuan Huang  <http://orcid.org/0000-0002-2215-2473>  
 Doyle V. Ward  <http://orcid.org/0000-0002-4952-824X>  
 Beth A. McCormick  <http://orcid.org/0000-0001-5992-2673>  
 John P. Haran  <http://orcid.org/0000-0001-7311-1121>

## References

- [1] Brubaker L, Putonti C, Dong Q, et al. The human urobiome. *Mamm Genome*. 2021;32(4):232–238. doi: [10.1007/s00335-021-09862-8](https://doi.org/10.1007/s00335-021-09862-8)
- [2] Chai TC, Wolfe AJ, Brubaker L. The urinary microbiome: improving diagnostics and management of urinary tract infections in adult females. *Infect Dis Clin North Am*. 2024;38(2):241–253. doi: [10.1016/j.idc.2024.03.003](https://doi.org/10.1016/j.idc.2024.03.003)
- [3] Hilt EE, McKinley K, Pearce MM, et al. Urine is not sterile: use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder. *J Clin Microbiol*. 2014;52(3):871–876. doi: [10.1128/JCM.02876-13](https://doi.org/10.1128/JCM.02876-13)
- [4] Mueller ER, Wolfe AJ, Brubaker L. Female urinary microbiota. *Curr Opin Urol*. 2017;27(3):282–286. doi: [10.1097/MOU.0000000000000396](https://doi.org/10.1097/MOU.0000000000000396)
- [5] Price TK, Hilt EE, Thomas-White K, et al. The urobiome of continent adult women: a cross-sectional study. *BJOG*. 2020;127(2):193–201. doi: [10.1111/1471-0528.15920](https://doi.org/10.1111/1471-0528.15920)
- [6] Price TK, Wolff B, Halverson T, et al. Temporal dynamics of the adult female lower urinary tract microbiota. *MBio*. 2020;11(2). doi: [10.1128/mBio.00475-20](https://doi.org/10.1128/mBio.00475-20)
- [7] Wolfe AJ, Brubaker L. Urobiome updates: advances in urinary microbiome research. *Nat Rev Urol*. 2019;16(2):73–74. doi: [10.1038/s41585-018-0127-5](https://doi.org/10.1038/s41585-018-0127-5)
- [8] Wolfe AJ, Toh E, Shibata N, et al. Evidence of uncultivated bacteria in the adult female bladder. *J Clin Microbiol*. 2012;50(4):1376–1383. doi: [10.1128/JCM.05852-11](https://doi.org/10.1128/JCM.05852-11)
- [9] Brubaker L, Wolfe AJ. The female urinary microbiota, urinary health and common urinary disorders. *Ann Transl Med*. 2017;5(2):34. doi: [10.21037/atm.2016.11.62](https://doi.org/10.21037/atm.2016.11.62)
- [10] Hourigan SK, Zhu W, Sw Wong W, et al. Studying the urine microbiome in superficial bladder cancer: samples obtained by midstream voiding versus cystoscopy. *BMC Urol*. 2020;20(1):5. doi: [10.1186/s12894-020-0576-z](https://doi.org/10.1186/s12894-020-0576-z)
- [11] Cortes-Penfield NW, Trautner BW, Jump RLP. Urinary tract infection and asymptomatic bacteriuria in older adults. *Infect Disease Clin NA*. 2017;31(4):673–688. doi: [10.1016/j.idc.2017.07.002](https://doi.org/10.1016/j.idc.2017.07.002)
- [12] Bonkat G, Cai T, Veeratterapillay R, et al. Management of urosepsis in 2018. *Eur Urol Focus*. 2019;5(1):5–9. doi: [10.1016/j.euf.2018.11.003](https://doi.org/10.1016/j.euf.2018.11.003)
- [13] Nicolle LE, Gupta K, Bradley SF, et al. Clinical practice guideline for the management of asymptomatic bacteriuria: 2019 update by the Infectious Diseases Society of America. *Clin Infect Dis*. 2019;68(10):e83–e110. doi: [10.1093/cid/ciy121](https://doi.org/10.1093/cid/ciy121)
- [14] Luu T, Albarillo FS. Asymptomatic bacteriuria: prevalence, diagnosis, management, and current antimicrobial stewardship implementations. *Am J Med*. 2022;135(8):e236–e44. doi: [10.1016/j.amjmed.2022.03.015](https://doi.org/10.1016/j.amjmed.2022.03.015)
- [15] Pallin DJ, Ronan C, Montazeri K, et al. Urinalysis in acute care of adults: pitfalls in testing and interpreting results. *Open Forum Infect Dis*. 2014;1(1):ofu019. doi: [10.1093/ofid/ofu019](https://doi.org/10.1093/ofid/ofu019)
- [16] McKenzie R, Stewart MT, Bellantoni MF, et al. Bacteriuria in individuals who become delirious. *Am J Med*. 2014;127(4):255–257. doi: [10.1016/j.amjmed.2013.10.016](https://doi.org/10.1016/j.amjmed.2013.10.016)
- [17] Nicolle L. Urinary tract infections in the older adult. *Clin Geriatr Med*. 2016;32(3):523–538. doi: [10.1016/j.cger.2016.03.002](https://doi.org/10.1016/j.cger.2016.03.002)
- [18] Petty LA, Vaughn VM, Flanders SA, et al. Assessment of testing and treatment of asymptomatic bacteriuria initiated in the emergency department. *Open Forum Infect Dis*. 2020;7(12):ofaa537. doi: [10.1093/ofid/ofaa537](https://doi.org/10.1093/ofid/ofaa537)
- [19] Rowe TA, Juthani-Mehta M. Diagnosis and management of urinary tract infection in older adults. *Infect Dis Clin North Am*. 2014;28(1):75. doi: [10.1016/j.idc.2013.10.004](https://doi.org/10.1016/j.idc.2013.10.004)
- [20] Ruth Kandel DP, Doron S. Do not test, do not treat asymptomatic bacteriuria. *Health MDOP*, editor. Massachusetts partnership for infection prevention. The Massachusetts Department of Public Health: Boston MA. 2009.
- [21] Hartley S, Valley S, Kuhn L, et al. Inappropriate testing for urinary tract infection in hospitalized patients: an opportunity for improvement. *Infect Control Hosp Epidemiol*. 2015;34(11):1204–1207. doi: [10.1086/673449](https://doi.org/10.1086/673449)
- [22] Yin P, Kiss A, Leis JA. Urinalysis orders among patients admitted to the general medicine service. *JAMA Intern Med*. 2015;175(10):1711–1713. doi: [10.1001/jamainternmed.2015.4036](https://doi.org/10.1001/jamainternmed.2015.4036)
- [23] Lloyd AL, Henderson TA, Vigil PD, et al. Genomic islands of uropathogenic Escherichia coli contribute to virulence. *J Bacteriol*. 2009;191(11):3469–3481. doi: [10.1128/JB.01717-08](https://doi.org/10.1128/JB.01717-08)
- [24] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–2120. doi: [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170)
- [25] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–359. doi: [10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923)
- [26] Blanco-Miguez A, Beghini F, Cumbo F, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat Biotechnol*. 2023;41(11):1633–1644. doi: [10.1038/s41587-023-01688-w](https://doi.org/10.1038/s41587-023-01688-w)
- [27] Beghini F, McIver LJ, Blanco-Miguez A, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with biobakery 3. *eLife*. 2021;10:10. doi: [10.7554/eLife.65088](https://doi.org/10.7554/eLife.65088)
- [28] Liu B, Zheng D, Zhou S, et al. Vfdb 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res*. 2022;50(D1):D912–D917. doi: [10.1093/nar/gkab1107](https://doi.org/10.1093/nar/gkab1107)
- [29] Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using diamond. *Nat Methods*. 2021;18(4):366–368. doi: [10.1038/s41592-021-01101-x](https://doi.org/10.1038/s41592-021-01101-x)
- [30] Oksanen J, Simpson GL, Blanchet FG, et al. Vegan: community ecology package. R package version 2.6–4. 2022. Available from: <https://CRAN.R-project.org/package=vegan>
- [31] McMurdie PJ, Holmes S, Watson M. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE*. 2013;8(4):e61217. doi: [10.1371/journal.pone.0061217](https://doi.org/10.1371/journal.pone.0061217)
- [32] Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics

- studies. *PLOS Comput Biol.* **2021**;17(11):e1009442. doi: [10.1371/journal.pcbi.1009442](https://doi.org/10.1371/journal.pcbi.1009442)
- [33] Wickham H. *Ggplot2: elegant graphics for data analysis.* (NY): Springer-Verlag; **2016**.
- [34] Namkung J. Machine learning methods for microbiome studies. *J Microbiol.* **2020**;58(3):206–216. doi: [10.1007/s12275-020-0066-8](https://doi.org/10.1007/s12275-020-0066-8)
- [35] Zhang Z, Lai Z, Xu Y, et al. Discriminative elastic-net regularized linear regression. *IEEE Trans Image Process.* **2017**;26(3):1466–1481. doi: [10.1109/TIP.2017.2651396](https://doi.org/10.1109/TIP.2017.2651396)
- [36] Ranstam J, Cook JA. Lasso regression. *Br J Surg.* **2018**;105(10):1348–1348. doi: [10.1002/bjs.10895](https://doi.org/10.1002/bjs.10895)
- [37] McDonald GC. Ridge regression. *WIREs Comput Stats.* **2009**;1(1):93–100. doi: [10.1002/wics.14](https://doi.org/10.1002/wics.14)
- [38] Brereton RG, Lloyd GR. Support vector machines for classification and regression. *Analyst.* **2010**;135(2):230–267. doi: [10.1039/B918972F](https://doi.org/10.1039/B918972F)
- [39] Breiman L. Random forests. *Mach Learn.* **2001**;45(1):5–32. doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)
- [40] Arik SÖ, Pfister T, editors. *Tabnet: attentive interpretable tabular learning.* Proceedings of the AAAI conference on artificial intelligence. Association for Computing Machinery, New York NY United States: San Francisco CA. **2021**.
- [41] Chen T, Guestrin C, editors. *Xgboost: a scalable tree boosting system.* Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. Palo Alto CA: AAAI Press. **2016**.
- [42] Li P, Luo J, Chen J, et al. The association between long-acting muscarinic antagonist-based therapy and the risk of urinary tract infection in patients with chronic obstructive pulmonary disease. *Eur J Intern Med.* **2024**;128:145–147. doi: [10.1016/j.ejim.2024.05.022](https://doi.org/10.1016/j.ejim.2024.05.022)
- [43] Sharafkhaneh A, Majid H, Gross NJ. Safety and tolerability of inhalational anticholinergics in COPD. *Drug Healthc Patient Saf.* **2013**;5:49–55. doi: [10.2147/DHPS.S7771](https://doi.org/10.2147/DHPS.S7771)
- [44] Bradley ES, Schell B, Ward DV, et al. The urinary microbiome of older adults residing in a nursing home varies with duration of residence and shows increases in potential pathogens. *J Gerontol A Biol Sci Med Sci.* **2022**;77(7):1313–1320. doi: [10.1093/gerona/glab345](https://doi.org/10.1093/gerona/glab345)
- [45] Gilbert JA, Blaser MJ, Caporaso JG, et al. Current understanding of the human microbiome. *Nat Med.* **2018**;24(4):392–400. doi: [10.1038/nm.4517](https://doi.org/10.1038/nm.4517)
- [46] Yu Y, Zielinski MD, Rolfe MA, et al. Similar neutrophil-driven inflammatory and antibacterial responses in elderly patients with symptomatic and asymptomatic bacteruria. *Infect Immun.* **2015**;83(10):4142–4153. doi: [10.1128/IAI.00745-15](https://doi.org/10.1128/IAI.00745-15)
- [47] Man SM. The clinical importance of emerging *Campylobacter* species. *Nat Rev Gastroenterol Hepatol.* **2011**;8(12):669–685. doi: [10.1038/nrgastro.2011.191](https://doi.org/10.1038/nrgastro.2011.191)
- [48] Williams GD, Burnham C-AD. Two cases of urinary tract infection caused by *Propionimicrobium lymphophilum*. *J Clin Microbiol.* **2015**;53(9):3077–3080. doi: [10.1128/JCM.00438-15](https://doi.org/10.1128/JCM.00438-15)
- [49] Ulger-Toprak N, Lawson PA, Summanen P, et al. *Peptoniphilus duerdennii* sp. nov. and *Peptoniphilus koenoneniae* sp. nov. isolated from human clinical specimens. *Int J Syst Evol Microbiol.* **2012**;62(Pt 10):2336–2341. doi: [10.1099/ijss.0.031997-0](https://doi.org/10.1099/ijss.0.031997-0)
- [50] Anderson MJ, Walsh DCI. Permanova, anosim, and the Mantel test in the face of heterogeneous dispersions: what null hypothesis are you testing? *Ecol Monogr.* **2013**;83(4):557–574. doi: [10.1890/12-2010.1](https://doi.org/10.1890/12-2010.1)
- [51] George SD, Van Gerwen OT, Dong C, et al. The role of *Prevotella* species in female genital tract infections. *Pathogens.* **2024**;13(5):364. doi: [10.3390/pathogens13050364](https://doi.org/10.3390/pathogens13050364)
- [52] Acuña-Amador L, Barloy-Hubler F. Porphyromonas spp. have an extensive host range in ill and healthy individuals and an unexpected environmental distribution: a systematic review and meta-analysis. *Anaerobe.* **2020**;66. doi: [10.1016/j.anaerobe.2020.102280](https://doi.org/10.1016/j.anaerobe.2020.102280)
- [53] Citron DM, Tyrrell KL, Goldstein EJC. *Peptoniphilus coxii* sp. nov. and *Peptoniphilus tyrelliae* sp. nov. isolated from human clinical infections. *Anaerobe.* **2012**;18(2):244–248. doi: [10.1016/j.anaerobe.2011.11.008](https://doi.org/10.1016/j.anaerobe.2011.11.008)
- [54] Bradley ES, Zeamer AL, Bucci V, et al. Oropharyngeal microbiome profiled at admission is predictive of the need for respiratory support among COVID-19 patients. *Front Microbiol.* **2022**;13. doi: [10.3389/fmicb.2022.1009440](https://doi.org/10.3389/fmicb.2022.1009440)
- [55] Haran JP, Bradley E, Zeamer AL, et al. Inflammation-type dysbiosis of the oral microbiome associates with the duration of COVID-19 symptoms and long-COVID. *JCI Insight.* **2021**;6(20). doi: [10.1172/jci.insight.152346](https://doi.org/10.1172/jci.insight.152346)
- [56] Choi J, Thänert R, Reske KA, et al. Gut microbiome correlates of recurrent urinary tract infection: a longitudinal, multi-center study. *eClinicalmedicine.* **2024**;71:71. doi: [10.1016/j.eclim.2024.102490](https://doi.org/10.1016/j.eclim.2024.102490)
- [57] Bućević PV, Šitum M, Chow C-E, et al. The urinary microbiome associated with bladder cancer. *Sci Rep.* **2018**;8(1). doi: [10.1038/s41598-018-29054-w](https://doi.org/10.1038/s41598-018-29054-w)
- [58] Holland B, Karr M, Delfino K, et al. The effect of the urinary and faecal microbiota on lower urinary tract symptoms measured by the international prostate symptom score: analysis utilising next-generation sequencing. *BJU Int.* **2020**;125(6):905–910. doi: [10.1111/bju.14972](https://doi.org/10.1111/bju.14972)
- [59] Ju T, Kong JY, Stothard P, et al. Defining the role of *Parasutterella*, a previously uncharacterized member of the core gut microbiota. *ISME J.* **2019**;13(6):1520–1534. doi: [10.1038/s41396-019-0364-5](https://doi.org/10.1038/s41396-019-0364-5)
- [60] Chan CCY, Lewis IA. Role of metabolism in uropathogenic *Escherichia coli*. *Trends Microbiol.* **2022**;30(12):1174–1204. doi: [10.1016/j.tim.2022.06.003](https://doi.org/10.1016/j.tim.2022.06.003)
- [61] Alteri CJ, Smith SN, Mobley HL, et al. Fitness of *Escherichia coli* during urinary tract infection requires gluconeogenesis and the TCA cycle. *PLOS Pathog.* **2009**;5(5):e1000448. doi: [10.1371/journal.ppat.1000448](https://doi.org/10.1371/journal.ppat.1000448)
- [62] Reitzer L, Zimmern P. Rapid growth and metabolism of uropathogenic *Escherichia coli* in relation to urine composition. *Clin Microbiol Rev.* **2019**;33(1). doi: [10.1128/CMR.00101-19](https://doi.org/10.1128/CMR.00101-19)
- [63] Shea Allyson E, Marzoa J, Himpsl Stephanie D, et al. *Escherichia coli* CFT073 fitness factors during urinary tract infection: identification using an ordered transposon library. *Appl Environ Microbiol.* **2020**;86(13):e00691–20. doi: [10.1128/AEM.00691-20](https://doi.org/10.1128/AEM.00691-20)
- [64] Bielecki P, Muthukumarasamy U, Eckweiler D, et al. In vivo mRNA profiling of uropathogenic *Escherichia coli* from diverse phylogroups reveals common and

- group-specific gene expression profiles. *MBio*. 2014;5(4):e01075–14. doi: [10.1128/mBio.01075-14](https://doi.org/10.1128/mBio.01075-14)
- [65] Jones SA, Jorgensen M, Chowdhury FZ, et al. Glycogen and maltose utilization by *Escherichia coli* O157: h7 in the mouse intestine. *Infect Immun*. 2008;76(6):2531–2540. doi: [10.1128/IAI.00096-08](https://doi.org/10.1128/IAI.00096-08)
- [66] Alteri CJ, Mobley HLT, Conway T, et al. Metabolism and fitness of urinary tract pathogens. *Microbiol Spectr*. 2015;3(3). doi: [10.1128/microbiolspec.MBP-0016-2015](https://doi.org/10.1128/microbiolspec.MBP-0016-2015)
- [67] Forsyth VS, Armbruster CE, Smith SN, et al. Rapid growth of uropathogenic *Escherichia coli* during human urinary tract infection. *MBio*. 2018;9(2). doi: [10.1128/mBio.00186-18](https://doi.org/10.1128/mBio.00186-18)
- [68] Wright KJ, Seed PC, Hultgren SJ. Uropathogenic *Escherichia coli* flagella aid in efficient urinary tract colonization. *Infect Immun*. 2005;73(11):7657–7668. doi: [10.1128/IAI.73.11.7657-7668.2005](https://doi.org/10.1128/IAI.73.11.7657-7668.2005)
- [69] Kulkarni R, Dhakal BK, Slechta ES, et al. Roles of putative type II secretion and type IV pilus systems in the virulence of uropathogenic *Escherichia coli*. *PLOS ONE*. 2009;4(3):e4752. doi: [10.1371/journal.pone.0004752](https://doi.org/10.1371/journal.pone.0004752)
- [70] Backert S, Meyer TF. Type IV secretion systems and their effectors in bacterial pathogenesis. *Curr Opin Microbiol*. 2006;9(2):207–217. doi: [10.1016/j.mib.2006.02.008](https://doi.org/10.1016/j.mib.2006.02.008)
- [71] Drengenes C, Wiker HG, Kalanathan T, et al. Laboratory contamination in airway microbiome studies. *BMC Microbiol*. 2019;19(1):187. doi: [10.1186/s12866-019-1560-1](https://doi.org/10.1186/s12866-019-1560-1)
- [72] Salter SJ, Cox MJ, Turek EM, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol*. 2014;12(1):1–12. doi: [10.1186/s12915-014-0087-z](https://doi.org/10.1186/s12915-014-0087-z)
- [73] Wang Y, LeCao KA. Managing batch effects in microbiome data. *Brief Bioinform*. 2020;21(6):1954–1970. doi: [10.1093/bib/bbz105](https://doi.org/10.1093/bib/bbz105)