Characterising urinary tract and clostridiodes difficile infections in the Southwest of England with the BNSSG systemwide data set

Brian Sullivan[1], Axel Montout[1], Ed Barker[1], Ranjeet Bhamber[1], Yujie Dai[2], Ashley Hammond[1], Philip Williams[3,7], Elizabeth Beech[4], Alastair Hay[1,5], Mahableshwar Albur[6], Andrew Dowsey[1]

1. Population Health Sciences , University of Bristol; 2. Engineering Mathematics, University of Bristol; 3. Department of Microbiology, University Hospitals Bristol and Weston NHS Foundation Trust 4. NHS England; 5. Concord Medical Centre; 6. Severn Infection Sciences, Southmead Hospital, North Bristol NHS Trust; 7. UK Health Security Agency

CONTACT INFORMATION

brian.sullivan@birstol.ac.uk

yujie.dai@bristol.ac.uk

andrew.dowsey@bristol.ac.uk

INTRODUCTION

Bacterial infections present many problems for patients and clinicians ranging from unpleasant to life-threatening. In partnership with Bristol, North Somerset, South Gloucestershire (BNSSG) Integrated Care Board (ICB), we have obtained two large regional datasets to describe the characteristics and risk factors **for catheter associated urinary tract infections (CAUTIs)** and **clostridioides difficile infections (CDIs)** in BNSSG and develop prediction models for risk assessment and to support optimal antibiotic prescribing and stewardship. These infections are locally problematic, may involve strains resistant to antimicrobial treatment, and in some cases lead to sepsis which can be fatal. This comprehensive dataset contains routinely collected health data covering nearly 1M patients in the ICB area across sources from primary and secondary care and pathology laboratories.

AIM

Curate and transform disparate data sources into a coherent patient level description of daily medical history over a 3-year period

Describe general characteristics of the CDI, UTI (focused on catheter associated), and control populations

Use logistic regression models to estimate risk factors for UTI and CDI infections to inform local clinicians

Develop prediction and time-series models focused on these risk factors and antimicrobial stewardship

MATERIAL & METHODS

The BNSSG systemwide data set includes adult patient data regarding:

Primary care data: Demographics, living status, comorbidities

Primary care Prescription Dispensations: Antibiotics, immunosuppressants, catheters, gastrointestinal and nutrition, hormones.

Secondary care data: ICD-10 diagnosis and OPCS procedure codes related to infections, chemotherapy, and gastrointestinal and urinary health.

Pathology data: Urinary and blood bacterial culture results, antimicrobial resistance tests, stool tests for C. Difficile infection, and virology test results.

Data are linked at the patient level and aligned to create timeline of medical events over the three-year data capture period. Datasets were analysed separately, with similar methods, using analytic pipelines developed in Python and R.
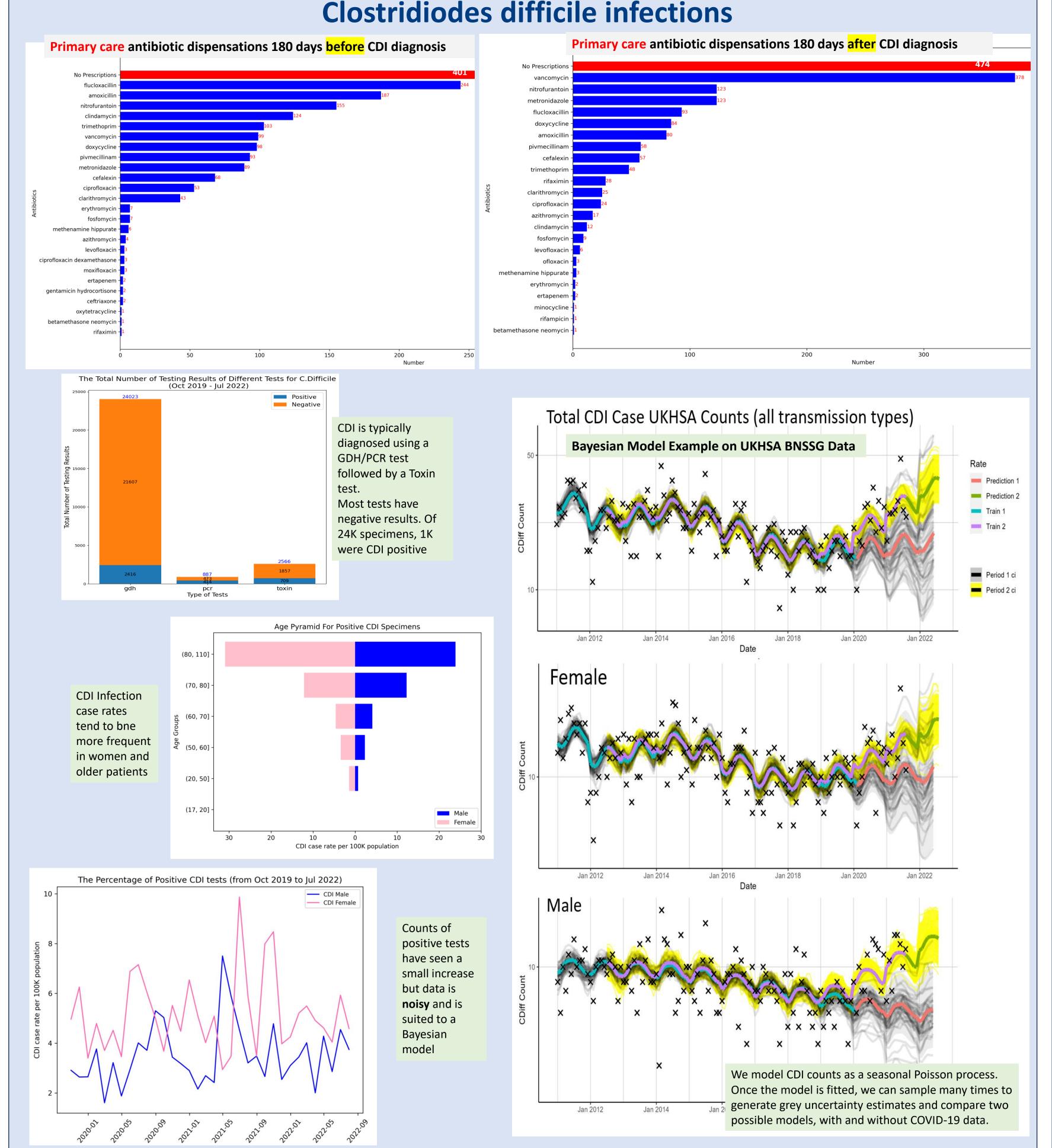
RESULTS

The dataset spans Oct 2019 to Jul 2022 and has records for 962K patients, with 7969 catheter patients (Indwelling: 5243, Intermittent: 2726)
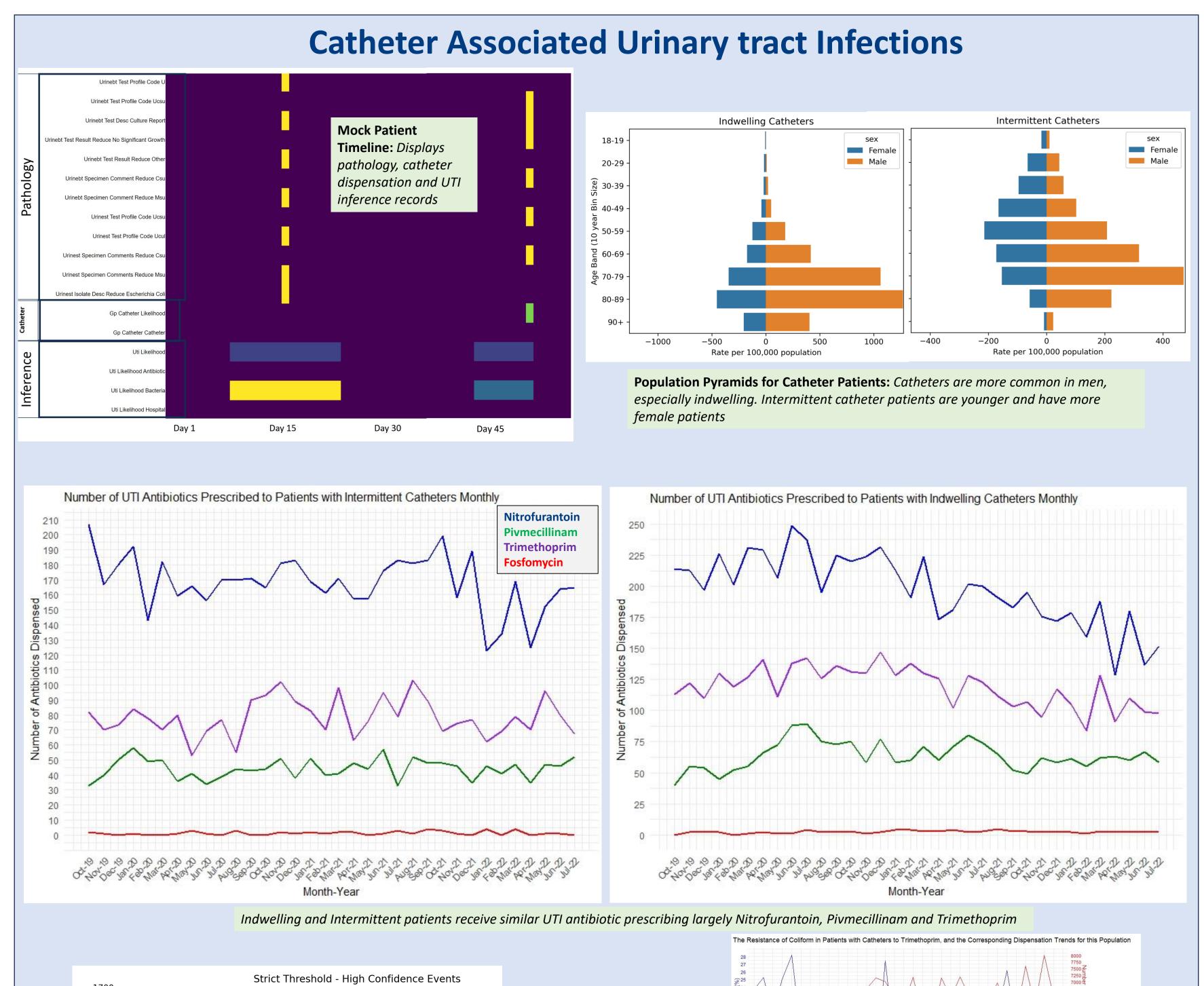
24K stool specimens were processed for CDI tests. 1091 patients had at least one positive CDI case (combination of positive GDH/PCR + Toxin).
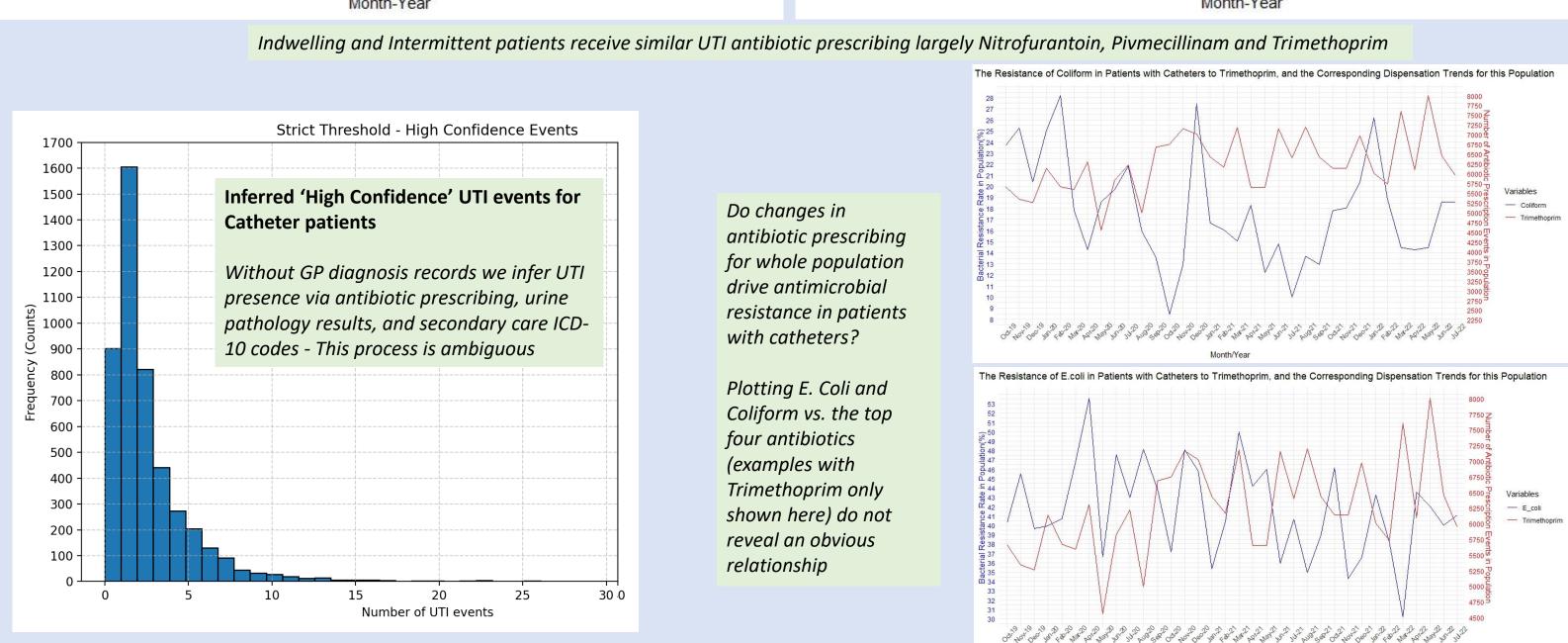
291K urine specimens were processed and 770K individual antimicrobial susceptibility tests (ASTs) were conducted of which 13% detected resistance, with Trimethoprim, Co-Amoxiclav, and Nitrofurantoin being the three most common. Escherichia coli, Proteus mirabilis, and Coliform spp were the three most common urine-borne bacteria with at least one antimicrobial resistance.

73K blood specimens were processed and 125K individual antimicrobial susceptibility tests (ASTs) were conducted of which 24% detected resistance, with Amp/Amoxicillin, Ampicillin, and Co-Amoxiclav being the most common. Escherichia coli, Klebsiella pneumoniae, and Methicillin Resistant Staphylococcus aureus were the three most common blood-borne bacteria with at least one antimicrobial resistance.

302K patients received at least one antibiotic prescription out of 912K AM dispensation events total.

Clostridiodes difficile infections

Primary care antibiotic dispensations 180 days **before** CDI diagnosis

Primary care antibiotic dispensations 180 days **after** CDI diagnosis

CDI is typically diagnosed using a GDH/PCR test followed by a Toxin test. Most tests have negative results. Of 24K specimens, 1K were CDI positive

CDI Infection case rates tend to bone more frequent in women and older patients

Counts of positive tests have seen a small increase but data is **noisy** and is suited to a Bayesian model

Bayesian Model Example on UKHSA BNSSG Data

We model CDI counts as a seasonal Poisson process. Once the model is fitted, we can sample many times to generate grey uncertainty estimates and compare two possible models, with and without COVID-19 data.

Catheter Associated Urinary tract Infections

Mock Patient Timeline: *Displays pathology, catheter dispensation and UTI inference records*

Population Pyramids for Catheter Patients: *Catheters are more common in men, especially indwelling. Intermittent catheter patients are younger and have more female patients*

Number of UTI Antibiotics Prescribed to Patients with Intermittent Catheters Monthly

Nitrofurantoin
Pivmecillinam
Trimethoprim
Fosfomycin

Number of UTI Antibiotics Prescribed to Patients with Indwelling Catheters Monthly

Indwelling and Intermittent patients receive similar UTI antibiotic prescribing largely Nitrofurantoin, Pivmecillinam and Trimethoprim

Inferred 'High Confidence' UTI events for Catheter patients

Without GP diagnosis records we infer UTI presence via antibiotic prescribing, urine pathology results, and secondary care ICD-10 codes – This process is ambiguous

Do changes in antibiotic prescribing for whole population drive antimicrobial resistance in patients with catheters?

Plotting E. Coli and Coliform vs. the top four antibiotics (examples with Trimethoprim only shown here) do not reveal an obvious relationship

SUMMARY / CONCLUSION

The BNSSG systemwide dataset offers an unprecedented view of linked patient level data at the ICB level over three years with day-to-day descriptions of medical events crucial to understanding bacterial infection, antimicrobial prescribing and resistance. We present a variety of visualisations and analyses of this comprehensive dataset. The project is a work in progress and given the scope of the dataset has many possible avenues for analysis and visualization.

We are currently in the descriptive phase of analysis but are preparing to create risk factor models. Such models require a great deal of clinical advice as the data often requires several decisions on how to transform into a format suitable for modelling (e.g., how to infer UTI, how to incorporate time-varying variables). Our initial UTI model will only consider catheter patients and predict if they did or did not have a UTI over the 3-year period. Demographics, living status, comorbidities, and catheter related parameters (type, size, material/coating) will be used as independent variables to predict the first UTI event recorded (if present).

More sophisticated models integrating time series information, antimicrobial resistance and other variables will be explored.