TITLE PAGE

FULL/LONG TITLE OF THE STUDY

Adaptive Prediction of Antimicrobial Susceptibility and its Implementation to Improve the Management of Urinary Tract Infection

SHORT STUDY TITLE / ACRONYM

ADAPT-AST (Adaptive Antimicrobial Susceptibility Testing)

PROTOCOL VERSION NUMBER AND DATE

1.3 October 2023

REFERENCE NUMBERS

SPONSORS Number: LHS0205

FUNDERS Number: 226691/Z/22/Z

IRAS Number: 330186

SIGNATURE PAGE

The undersigned confirm that the following protocol has been agreed and accepted and that the Chief Investigator agrees to conduct the study in compliance with the approved protocol and will adhere to the principles outlined in the Declaration of Helsinki, the Sponsor's SOPs, and other regulatory requirement.

I agree to ensure that the confidential information contained in this document will not be used for any other purpose other than the evaluation or conduct of the investigation without the prior written consent of the Sponsor

I also confirm that I will make the findings of the study publicly available through publication or other dissemination tools without any unnecessary delay and that an honest accurate and transparent account of the study will be given; and that any discrepancies from the study as planned in this protocol will be explained.

For and on behalf of the Study Sponsor:

Signature: Date:

...../...../.....

.....

Name (please print):

Position:

Chief Investigator:

Signature: Date:

...../...../.....

.....

Name: (please print):

LIST OF CONTENTS

LIST OF ABBREVIATIONS

AMR	Antimicrobial Resistance
APR	Annual Progress Report
AST	Antimicrobial Susceptibility Testing
AUC-ROC	Area Under the Curve – Receiver Operator Characteristic

AWaRe	Access, Watch, Reserve
BCI	Bayesian causal inference
BL	Beta-Lactam
BL-BLIC	Beta-Lactam-Beta-Lactamase-Inhibitor Combination
CAMO-NET	Centres for Antimicrobial Optimisation Network
CI	Chief Investigator
CIPHA	Combined Intelligence for Public Health Action
CDC	Civic Data Co-operative
COVID-19	Coronavirus Disease 2019
CQUIN	Care Quality Indicator
CTIMP	Clinical Trial of an Investigational Medicinal Product
DAAG	Data Asset and Access Group
DARG	Data Action Research Group
DSA	Data Sharing Agreement
DPIA	Data Protection Impact Assessment
eGFR	Estimated Glomerular Filtration Rate
EPMA	Electronic Prescribing and Medicines Administration
EMIS	Egton Medical Information Systems
EUCAST	European Committee on Antimicrobial Susceptibility Testing
GP	General Practice
IV	Intravenous
LCL	Liverpool Clinical Laboratories
LCH	Liverpool Community Health
LHP	Liverpool Health Partners
LMIC	Low or Middle Income
LR	Logistic Regression
LUHFT	Liverpool University Hospitals NHS Foundation Trust
LWfT	Liverpool Womens' NHS Foundation Trust
MASCC	Multinational Association of Supportive Care in Cancer
MCMC	Markov Chain Monte Carlo
NHS	National Health Service

NIHR	National Institute for Health Research
OPAT	Outpatient Parenteral Antimicrobial Therapy
R&D	Research and Development
SNOMED-CT	Systematized Nomenclature of Medicine Clinical Terms
SMG	Study Management Group
SOP	Standard Operating Procedure
TRE	Trusted Research Environment
UKRI	United Kingdom Research and Innovation
UTI	Urinary Tract Infection
WHO	World Health Organisation

KEY STUDY CONTACTS

Chief Investigator	Alex Howard Medical Microbiology Department of Infection and Immunity Liverpool Clinical Laboratories Clinical Services Support Building Liverpool University Hospitals NHS Foundation Trust United Kingdom +44 7855 550 575 Alexander.Howard@liverpool.ac.uk
Sponsor	Liverpool University Hospitals NHS Foundation Trust is the Sponsor for this Study. It is recognised that as an employee of the Trust the Chief Investigator has been delegated specific duties, as detailed in the Sponsorship Approval letter.
Funder(s)	Wellcome Trust Gibbs Building 215 Euston Road London NW1 2BE 020 7611 8888
Key Protocol Contributors	Dr Alex Howard Prof. William Hope Prof. Iain Buchan Dr David Hughes Prof. Simon Maskell Dr Beth Woods Liverpool AMR Citizens' Jury

STUDY SUMMARY

The aim of this study is to develop and evaluate ADAPT-AST, an adaptive informatics approach for laboratory antimicrobial susceptibility testing (AST) for urinary tract infection (UTI) pathogens compared with current practice, as a potential means to improve patient outcomes, reduce AMR risks and reduce waste of laboratory resources.

UTI is a leading cause of community and hospital acquired infection and a major driver of antimicrobial prescribing in primary and secondary care, exposing patients to the associated risks of antimicrobial resistance (AMR), drug toxicity and other healthcare-associated infections. The continued proliferation of AMR also increasingly limits treatment choices for many UTIs.

Despite the importance of UTI, antimicrobial susceptibility testing (AST) of urine specimens is based on inflexible 'one-size-fits' all standard operating procedures (SOPs). Either a very large unfocused panel of antimicrobials is immediately tested (leading to wasted resources), or more commonly, and particularly in low or middle income (LMIC) settings, a selected subset of antimicrobials is tested at day one prior to a second or even third panel of antimicrobials. Such an approach does not adapt to prior information such as previous resistance patterns, antimicrobial prescribing, or demographic information, despite these factors being powerful (strong) predictors of resistance. This results in imprecise, inefficient, and inequitable provision of antimicrobial susceptibility information, which provides suboptimal support of decisions for treatment of UTI.

This project will train adaptive algorithms (including Bayesian causal inference [BCI] algorithms) to predict urine AST results and prioritise testing using patient demographics, prescribing, admission, and microbiology laboratory care data. The clinical utility of these algorithms will be evaluated in terms of their ability to increase the number, timeliness and appropriateness of usable AST results available to clinicians, and their ability to reduce laboratory resource costs through better test prioritisation. The anticipated benefits of a successfully developed, evaluated, and implemented system are faster and more precise treatments of UTI in patients with drug-resistant organisms and more efficient resource management, particularly in laboratory and pharmacy workflows.

Study Title	Adaptive Prediction of Antimicrobial Susceptibility and its Implementation to Improve the Management of Urinary Tract Infection
Internal ref. no. (or short title)	ADAPT-AST (Adaptive Antimicrobial Susceptibility Testing)

Study Design	Retrospective observational study
Study Participants	Demographic, antimicrobial prescribing and microbiology specimen data from patients >18 years in the catchment area served by Liverpool University Hospitals NHS Foundation Trust and Liverpool Community Health
Planned Size of Sample (if applicable)	> 250,000 specimens estimated (the number of patients will be fewer because some patients will have had more than one specimen sent)
Outcomes	Primary outcome: The number of antimicrobial susceptibility results that 'should' spur action by a clinician on the day that actual first-line AST results were reported (in green on Figure 5). Secondary outcomes: 1. The number of days until a result that 'should' spur action by a clinician (in green on Figure 5). 2. The number of antimicrobial susceptibility results that 'could' spur action by a clinician on the day that actual first-line AST results were reported (in yellow on Figure 5). 3. The number of days until a result that 'could' spur action by a clinician (in yellow on Figure 5). 4. The projected health economic cost per specimen, including laboratory (e.g., consumable cost) and patient (e.g., drug toxicity, clinical failure) measures guided by the above criteria.
Follow up duration (if applicable)	Not applicable
Planned Study Period	3 years
Project Question/Aim(s)	1. Can causal, explainable Bayesian causal inference (BCI) approaches be used to develop/(re)validate (train/test) clinical prediction models for prediction of urine organism antimicrobial susceptibility using NHS care data?

	2. Can the resulting ADAPT-AST adaptive, informatics-based (data-driven and system-wide) approach target AST for UTI better than existing approaches can, thereby improving timeliness of clinical decision making and saving resources through efficiency?

FUNDING AND SUPPORT IN KIND

FUNDER(S)	FINANCIAL AND NON FINANCIALSUPPORT GIVEN
Wellcome Trust	Funding of all materials and staffing to conduct study

ROLE OF STUDY SPONSOR AND FUNDER

Study Sponsor (Liverpool University Hospitals NHS Foundation Trust):

The sponsor will assume overall responsibility for proportionate, effective arrangements being in place to set up, run and report the project. They will have overall responsibility for the project, including:

1. Identifying and addressing problems with the proposal, protocol and applications and ensuring that they take into account systematic reviews of relevant existing research evidence and other relevant projects in progress, make appropriate use of patient, service user and public involvement, and are scientifically sound (e.g. through independent expert review), safe, ethical, legal and feasible and remain so for the duration of the project, taking account of developments while the project is ongoing
2. Satisfying itself that the investigators, project team and project sites are suitable
3. Ensuring that roles and responsibilities of the parties involved in the project and any delegation by the sponsor of its tasks are agreed and documented
4. Ensuring adequate provision is made for insurance or indemnity to cover liabilities which may arise in relation to the design, management and conduct of the project
5. Ensuring appropriate arrangements are made for making information about the project publicly available before it starts

6. Agreeing appropriate arrangements for making data and tissue accessible, with adequate consent and privacy safeguards, in a timely manner after it has finished
7. Ensuring arrangements for information about the findings of the project to be made available, including, where appropriate, to participants)
8. Ensuring that, where expected or required, the project has approval from a research ethics committee (Whether outright or following a provisional opinion, re-submission, or appeal) and any other relevant approval bodies before it begins
9. Verifying that regulatory and practical arrangements are in place, before permitting the project to begin in a safe and timely manner
10. Putting and keeping in place arrangements for adequate finance and management of the project, including its competent risk management and data management
11. Ensuring that effective procedures and arrangements are kept in place and adhered to for reporting (e.g. progress reports, safety reports) and for monitoring the project, including its conduct and the ongoing suitability of the approved proposal or protocol considering adverse events or other developments.

Study funder (Wellcome Trust):

The role of the Wellcome Trust as study funder is limited to provision of funds to conduct the project. It has no role in the design or implementation of the study. Grant reference number 226691/Z/22/Z.

PROTOCOL CONTRIBUTORS

Dr Alex Howard (Chief Investigator) – writing

Prof. William Hope (Consultant in Infectious Diseases/Medical Microbiology) – review and editing

Prof. Iain Buchan (Consultant in Public Health) – review and editing

Dr David Hughes (Lecturer in Health Data Science) – review and editing

Prof. Simon Maskell (Professor of Autonomous Systems) – review and editing

Liverpool AMR Citizens' Jury – Patient participation and Involvement in general principles of project design (see sections 1.3 and 4.4)

Dr Beth Woods (Lecturer in Health Economics) – review and editing

Prof Mark Sculpher (Lecturer in Health Economics) – review and editing

Prof Marta Soares (Lecturer in Health Economics) – review and editing

KEY WORDS: Urinary tract infection, Antimicrobial Resistance, Susceptibility testing, Microbiology, Bayesian causal inference

STUDY PLAN

STUDY PROTOCOL

1 BACKGROUND and RATIONALE

The aim of this study is to develop and evaluate ADAPT-AST, an adaptive informatics (data-driven, whole-system) approach to improve decision-making over laboratory antimicrobial susceptibility testing (AST) for urinary tract infection (UTI) pathogens compared with current practice, and thereby examine the potential to improve patient outcomes, reduce AMR risks and reduce waste of laboratory resources.

1.1 Urinary Tract Infection and Current Laboratory Standard Operating Procedures

Urinary tract infection (UTI) is one of the leading causes of serious community and hospital acquired infection(1,2). UTI is therefore a major driver of antimicrobial prescribing in primary and secondary care, with inherent risks of antimicrobial resistance (AMR), drug toxicity and other healthcare-associated infections. The proliferation of AMR is also limiting treatment choices for many UTIs(3). Despite this, antimicrobial susceptibility testing (AST) of urine specimens is based on inflexible Standard Operating Procedures (SOPs) that are not fit for purpose in the era of AMR (see Figures 1 and 2).

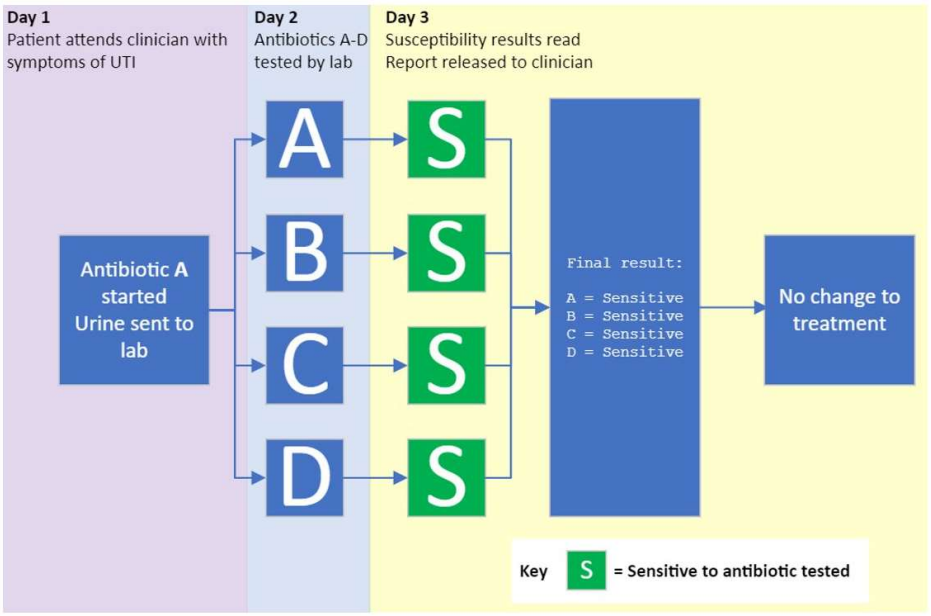

Figure 1: Current laboratory workflows fail to make use of past information related to the patient or the patient's immediate environment to make decisions about susceptibility testing, which may lead to delays in appropriate decision making. In low prevalence AMR settings, testing contributes little to clinical care.

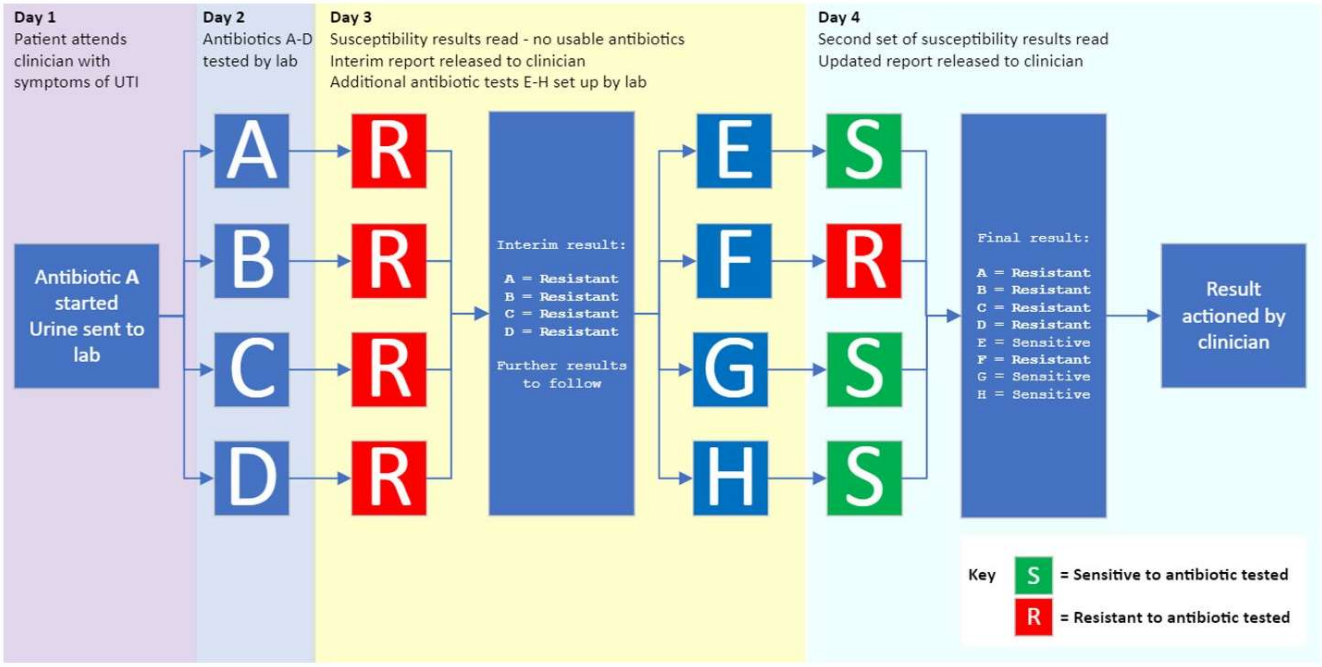

Figure 2: In highly-prelavent AMR settings, a second round of susceptibility testing is required before appropriate antimicrobials are identified and therapy is changed. Each additional panel of susceptibility testing typically incurs at least 24 hours of delay.

This standard practice results in the provision of susceptibility information for clinicians which is:

a) Imprecise

Some laboratories in high income settings can test very large panels of antimicrobials immediately at day one (for example, the ThermoFisher® Sensititre™ Gram Negative MIC Plate tests 22 antimicrobials), but more commonly, and particularly in low or middle income (LMIC) settings, a selected subset of antimicrobials is tested at day one to save resources (for example, the MAST URI® system used in Liverpool Clinical Laboratories [LCL] tests a 7-antimicrobial panel) (4–6). There is currently a 'one-size-fits-all' approach to this testing, where a standard panel of antimicrobials are initially tested against urinary pathogens. These panels do not adapt for previous resistance patterns, antimicrobial prescribing, or demographic information despite their potential impact on the likelihood of susceptibility to various agents(7).

b) Inefficient

When the standard panel tested yields no appropriate antibiotics for the clinician to use, a 'second line' standard panel of antibiotics will then be tested. This process unnecessarily consumes the time of laboratory staff and requires additional resources to those used to set up the initial unusable panel of antimicrobials. Lack of information related to appropriate oral or outpatient parenteral antimicrobial therapy (OPAT) treatment options delays early discharge with patients required to stay in hospital to receive high-frequency intravenous (IV) antibiotic regimens(8).

c) Inequitable

Current approaches of trial-and-error lead to a delay in actionable results being available to the clinician, and therefore a delay in the patient being changed to the appropriate antimicrobial therapy. The outcome of this may range from unnecessary adverse effects from the initial ineffective agent to progression of infection causing admission to hospital with sepsis(9). These complications are likely to disproportionately affect those at the highest risk of AMR, in whom the consequences of treatment failure are also often the most severe(10).

1.2 Bayesian causal inference, Prediction and Urinary Tract Infection

Predictive statistical modelling is a key pathway to the achievement of 'precision' medicine, i.e., medical management tailored to an individual's specific circumstances. Clinical prediction tools (for example, CURB-65 and MASCC [Multinational Association of Supportive Care in Cancer] scores in

pneumonia and neutropenic fever respectively) have traditionally been limited to crude risk stratification of patients into groups based on statistical analysis of evidence harvested from pre-existing observational studies(11,12). The quality and quantity of data available in clinical studies with relatively strict inclusion and exclusion criteria is often relatively low; this results in clinical tools that do not generalise well and cannot reliably be applied outside the study population in which they have been developed(13,14).

Over the past decade, three main developments have increased the potential power of clinical prediction tools. Firstly, the proliferation of electronic health records in clinical care has significantly increased the volume of structured clinical data that is potentially accessible for statistical analyses(15). Secondly relevant statistical and computational method/resources, and their ease of application to clinical data/workflows, have advanced(16). Thirdly, adaptive approaches to clinical prediction have been developed through a combination of biostatistical, Bayesian causal inference (BCI) and clinical informatics research(17). These methods build on traditional statistical approaches, allowing a flexibility of methodology that can improve clinical prediction as core business of health systems(18).

Infection management has been of particular interest in the field of precision medicine because the threat of sepsis combined with the ongoing proliferation of antimicrobial resistance means that antibiotic therapy needs to be not only efficacious, but also prompt and targeted(19). Several studies have employed statistical techniques to predict antimicrobial susceptibility testing in UTI. Methods for these applications have included logistic regression (LR), decision trees, random forest models, neural networks, and extreme gradient boosting. The predictive performance of models versus eventual culture(s) and clinical outcomes are evaluated using metrics such as area under the curve receiver operating characteristic (AUC-ROC). The objective of these studies, however, has been very different to that of ADAPT-AST; this prior work has focused solely on clinical decision support for clinician antimicrobial treatment selection at the bedside without consideration of the role of the laboratory. No studies to date have used adaptive prediction techniques to help personalise AST(20–24).

The findings of previous attempts to predict antimicrobial susceptibility algorithmically have been promising. The value of a bedside clinical decision tool, however, is affected by what has been referred to as the ‘chasm’ between demonstrating algorithmic accuracy and clinical effectiveness. This is due to a range of factors, which can only be mitigated by deep a priori knowledge of the clinical subject matter; for example, predictions based on AMR datasets are particularly vulnerable to ‘concept shift’ of training datasets caused by changes in organism ecology and clinical laboratory processes over time(25). Prescriptively directing antimicrobial administration using predictive algorithms is also problematic in the clinical setting where there is inherent situational uncertainty and the impact of

giving the wrong drug can be very high(26). Existing studies with limited specialist clinical involvement also suffer from an inability to separate causation from prediction, a particularly key distinction in medical microbiology where the complex host-microbe dynamic means that causal inference is not always straightforward(27).

In the laboratory the potential value of BCI in automating and prioritising specimen workflow has been recognised; for instance, the potential of adaptive informatics systems to reduce workload by avoiding unnecessary urine culture has been explored(28,29). There has, however, been very little adaptive prediction modelling utilising medical microbiology's unique position at the interface between the laboratory and the bedside; guiding diagnostic stewardship using methods which recognise and quantify inherent prediction uncertainty could provide a more tangible route to improving clinician decision making and driving resource efficiency in the real world(30).

1.3 ADAPT-AST: An Adaptive Antimicrobial Susceptibility Testing Framework

This study will develop and evaluate ADAPT-AST, an adaptive AST system trained on healthcare data from patients over the age of 18 under the care of Liverpool University Hospitals NHS Foundation Trust (LUHFT) and/or Mersey Care NHS Foundation Trust Primary Care facilities in the Liverpool area. The study will advance understanding of how adaptive informatics approaches can be used to better target AST in microbiology laboratories, regardless of the level of resource available. If implementable, the immediate potential benefits would be faster appropriate treatment of UTI in patients with drug-resistant organisms, and better resource utilisation. The system developed from these models would be much more relevant to AMR, targeted and implementable than those developed by previous adaptive informatics work in UTI, which has predominantly examined support of decisions over whether to culture and which antimicrobial agent to administer.

The project will examine how current causal BCI approaches perform this adaptive AST system. In a time of unprecedented resource pressures globally on healthcare and ever-increasing AMR, ADAPT-AST could significantly increase the efficiency of urine testing and quality of UTI care in a range of healthcare settings. A revised informatics-based strategy would align with both local and global healthcare policy objectives, reflected by the UK Government's 5-year National AMR Plan, NHS England's 2022/23 Care Quality Indicator (CQUIN) on the diagnosis and management of UTI, NHS Standard Contract obligations on antimicrobial use aligned to the World Health Organisation's (WHO's) Access, Watch Reserve (AWaRe) classification, and the UK Government's Carter Report on productivity in NHS hospitals(31–35). In 2022, an AMR Citizens' Jury was commissioned by the

University of Liverpool to explore public and patient perspectives on collaborative monitoring of AMR data for applications of this kind in the Liverpool area. The jury was broadly supportive of the use of integrated regional pseudonymised data for this process and raised important points about the legal, security and quality frameworks that will be required to underpin projects of this kind(36).

2 PROJECT QUESTIONS/AIMS

The aim of this project is to develop and evaluate adaptive informatics approaches to supporting better clinical decisions over laboratory antimicrobial susceptibility testing for UTI pathogens compared with current practice, and thereby examine the potential to improve patient outcomes, reduce AMR risks and reduce waste of laboratory resources. The project questions are as follows:

1. Can causal, explainable BCI approaches be used to develop/(re)validate (train/test) clinical prediction models for prediction of urine organism antimicrobial susceptibility using NHS care data?
2. Can the resulting ADAPT-AST adaptive, informatics-based (data-driven and system-wide) approach target AST for UTI better than existing approaches can, thereby improving timeliness of clinical decision making (reducing unnecessary AMR selection pressure) and saving resources through efficiency?

3 METHODS and THEORETICAL FRAMEWORK

The proposed approach is divided into 3 phases:

3.1 Data Collection and Integration

3.1.1 Setting and Eligibility Criteria

LUHFT and/or General Practices (GP) in the area served by Liverpool Clinical Laboratories have been chosen as the study setting to reflect the most representative range of patient types in the Liverpool area that data can become available for within the time frame of the study; other local Trusts have not

been included because they have many different clinical and laboratory information systems that would pose insurmountable challenges to extract information in a timely manner. The eligibility criteria for the study are:

- Inclusion criteria:
    - The specimens for which AST predictions & recommendations will be made are urine specimens processed by LCL Microbiology laboratory taken from patients ≥ 18 years old in LUHFT and/or GP locations that grew organisms within the period of the study dataset; these are the only specimens for which AST results will be available to train and test ADAPT-AST. Predictions will be made for all urine specimen types, including mid-stream urines, catheter specimens of urine and nephrostomy urine.
- Exclusion criteria:
    - Urine specimens processed by LCL that did not grow organisms within the period of the study dataset
    - Urine specimens taken from patients < 18 years old. These will be excluded because secondary care data is processed by Alder Hey Children's Hospital and will therefore be missing for this cohort
    - Predictions will be made for asymptomatic bacteriuria screening specimens in pregnant women who have had specimens sent from a GP, but not those which have been sent from Liverpool Womens' NHS Foundation Trust (LWfT). The reason for this is that some specimens sent from LWH will be from inpatients, for whom incomplete potentially important prescribing information close to the time point of sampling will be missing due to this data not being integrated.
    - Predictions for non-bacterial organisms grown in urine (i.e., fungi) will not be made; the small amount of susceptibility data available and the paucity of available agents will result in poorly-predictive algorithms with relatively little clinical impact.

3.1.2 Data Parameters of Interest

The pseudonymised data parameters of interest (see section 3.1.4) are listed in Table 1.

Data Category	Data Type
Demographic	Pseudonymised study number, pseudonymised postcode area, age group, sex

Admission / Consultation	Admission date & time, inpatient transfer date & time, pseudonymised inpatient transfer location type, discharge date and time, discharge location type, vital status (alive/deceased), date of change in vital status, general practitioner consultation date/time, previous general practitioner coded diagnoses
Prescribing	Drug name, pseudonymised prescription location type, route of administration, dose, dosing frequency, prescription start date & time, dates & times of administration, drug allergy (drug), drug allergy (reaction type), allergic reaction date
Microbiology / Pathology	Microbiology: Date & time of specimen collection, date & time of specimen receipt, data & time of specimen authorisation, specimen type, report code, comment code, specimen body site, pseudonymised location specimen was taken, pseudonymised specimen number, epithelial cells (urine specimens), white cell count (urine specimens), red cell count (urine specimens), organism count (urine specimens), organism code, organism name, antimicrobial susceptibilities for: phenoxymethylpenicillin, amoxicillin, ampicillin, oxacillin, methicillin, mecillinam, amoxicillin-clavulanate, piperacillin-tazobactam, temocillin, cefalexin, cefuroxime, cefpodoxime, ceftriaxone, cefoxitin, cefotaxime, ceftazidime, ceftazidime-avibactam, ceftolozane-tazobactam, ciprofloxacin, levofloxacin, gentamicin, amikacin, tetracycline, tigecycline, meropenem, ertapenem, teicoplanin, vancomycin, clindamycin, erythromycin, aztreonam, linezolid, co-trimoxazole, trimethoprim, nitrofurantoin, fosfomycin (IV), fosfomycin (oral), mupirocin, daptomycin, rifampicin, colistimethate sodium, metronidazole, chloramphenicol, fusidic acid. Biochemistry: Date of specimen collection, serum urea, serum creatinine, acute kidney injury alert flag

Table 1: Data that will be collected and assessed for inclusion in the predictive model.

The justifications for inclusion of the above data types are:

- Demographic data: Demographic features such as city area and sex may indirectly influence probability of antimicrobial resistance through factors such as socioeconomic status(37)
- Admission/consultation/past diagnoses data: Resistant organism colonisation is more likely in patients who have been exposed to healthcare(38).
- Prescribing data: Previous antimicrobial exposure is one of the key routes to development of drug resistance through induction of organism-level molecular resistance mechanisms and population-level selection of resistant organisms. Data on non-antimicrobial drugs will also be collected as this may be a surrogate marker of contact with healthcare that influences AMR risk e.g., drugs for benign prostatic hypertrophy(39)
- Microbiology data: Previous colonisation with an antimicrobial resistant organism is a risk factor for subsequent infection with an antimicrobial resistant organism; susceptibility data for all types of patient specimen will therefore be analysed (with appropriate algorithmic weighting), not just urine specimens(40). Specimen type and location type will also provide important information about risk factors (e.g., catheter specimen of urine, urology clinic specimens).

Protected characteristics were also considered for inclusion to prevent unintentional algorithmic discrimination. Characteristics considered were those defined by UK law: pregnancy, marital status, race, sexual orientation, gender reassignment, disability, religion and nationality (41,42). Data on these characteristics will not be collected because: 1. There are no proven or plausible reasons why these characteristics should directly and significantly impact AMR risk; 2. protected characteristic data contained within the relevant systems is likely to be incomplete and unreliable; 3. some protected characteristics (e.g., race or nationality) may increase the risk of participant identification due to the rarity of some characteristic types in the Liverpool area.

3.1.3 Sample size

The main sample size consideration is the volume of data required to train the BCI algorithm; this will be higher than the sample size required to provide 80% power to detect a significant (e.g., $p < 0.05$) difference between groups in the simulated interventional study. For the binary LR techniques like that used to develop the predictive model, however, much larger sample sizes are required to provide predictive accuracy(43). Given that a range of BCI techniques are likely to be trialled and that most of the variation will be taking place in a small proportion of samples, data pertaining to a large number of specimens will be required. All eligible patients who have grown an organism in a urine specimen during the a five year period from 17th May 2017 will be recruited for the development and testing of the predictive algorithm. This period has been chosen to provide a balance between strength of model, the amount of data available, computational power required and 'concept drift' i.e., changes in susceptibility patterns and laboratory practice over time; the most recent AST practice change in LCL was implemented 17th May 2017(44). This period will provide an estimated sample size of at least 250,000 specimens, which will pertain to fewer patients than that number as some will have had urine specimens sent more than once.

3.1.4 Pseudonymised Data Access

Patients do not exist solely within a single healthcare setting, but rather move through local healthcare systems according to their needs(45). The predictive value of the model will therefore theoretically be significantly improved by access to integrated data sets from across Liverpool primary and secondary care systems, to give an accurate picture of their care up to the point a urine sample is sent to the laboratory. This will be achieved using infrastructure provided by Mersey Care NHS Foundation

Trust's NHS Combined Intelligence for Population Health Action (CIPHA) programme supported by Liverpool City Region Civic Data Cooperative (CDC) and the CIPHA's Data Action Research Group (DARG), which are entities based within NHS Cheshire and Merseyside Integrated Care board (ICB) (46,47). Following relevant approvals (see section 4.2), researchers will be given access to data within the NHS ICB/CIPHA regional secure data environment (SDE) that has been pseudonymised to a level that prevents re-identification using any means reasonably likely to be used by either researchers or any other person, according to NHS Data Services for Commission Regional Offices (DSCRO) standard operating procedures (48):

- Removal of patient name and hospital number
- Replacement of NHS number with generic study number
- Replacement of specimen number with generic study number
- Replacement of locations with dummy locations and location type (e.g. medical ward, surgical ward, GP practice), removal of patient address and postcode pseudonymisation

Pseudonymisation code keys are kept within an encrypted drive within the NHS ICB/CIPHA secure data environment for 10 years from the end of the study, and are only accessible by ICB/CIPHA employees (they are not accessible by the research team).

Access to linked, pseudonymised data is the optimum approach for two reasons:

- Regional integration of hospital and community data is critical to this project. CIPHA is an entity designed with that sole purpose that has already proven itself effective in the COVID-19 pandemic(49). Its position as a link between different data controllers to generate social license for integrated data uses/intelligence is well placed to integrate hospital and community data into single pseudonymised datasets.
- LUHFT already has access to the necessary data science resource from University of Liverpool without depending on stretched Trust financial and human resource.

3.1.5 Data Cleaning

Once researchers have been given access to the data, cleaning of the dataset will be required because:

- Intrinsic resistance patterns will be populated in the dataset according to European Committee on Antimicrobial Susceptibility Testing (EUCAST) rules to enable proper interpretation (for

instance, amoxicillin will not be tested on most *Klebsiella* species because of intrinsic resistance, but this will be a blank in the starting dataset, which will need to be filled in with an 'R')(50)

- Organisms susceptible at increased antimicrobial exposure 'I' can be interpreted as susceptible in the context of urinary tract infection according to EUCAST rules(51). They will therefore be reclassified as 'S' to facilitate the use of binary LR (see below).
- Resistance data from multiple organisms grown in the same specimen will be inputted in the same way as for individual growth, but organisms will remain linked by pseudonymised specimen number to facilitate trial of composite 'functional' resistance of multiple organism growth as an input (i.e., if either organism in the specimen is resistant report the specimen as 'R', if both are susceptible report as 'S').
- Some data elements (for instance specimen site) will have been recorded manually and there may therefore be typographical errors that require correction to make them statistically processable.
- Assessment of data completeness will be required and strategies for managing missing data developed.
- Other data transformations may need to occur to facilitate processability or economise on computing power, e.g., feature scaling or dimensionality reduction(52).

3.2 Development of an Adaptive Clinical Prediction Model

3.2.1 Bayesian causal inference Model Training

All data manipulation, analysis and algorithm writing will take place within the CIPHA TRE. The output of the adaptive clinical prediction model will be a quantification of the level of statistical probability of susceptibility to a specified antibiotic, as displayed in Figure 3.

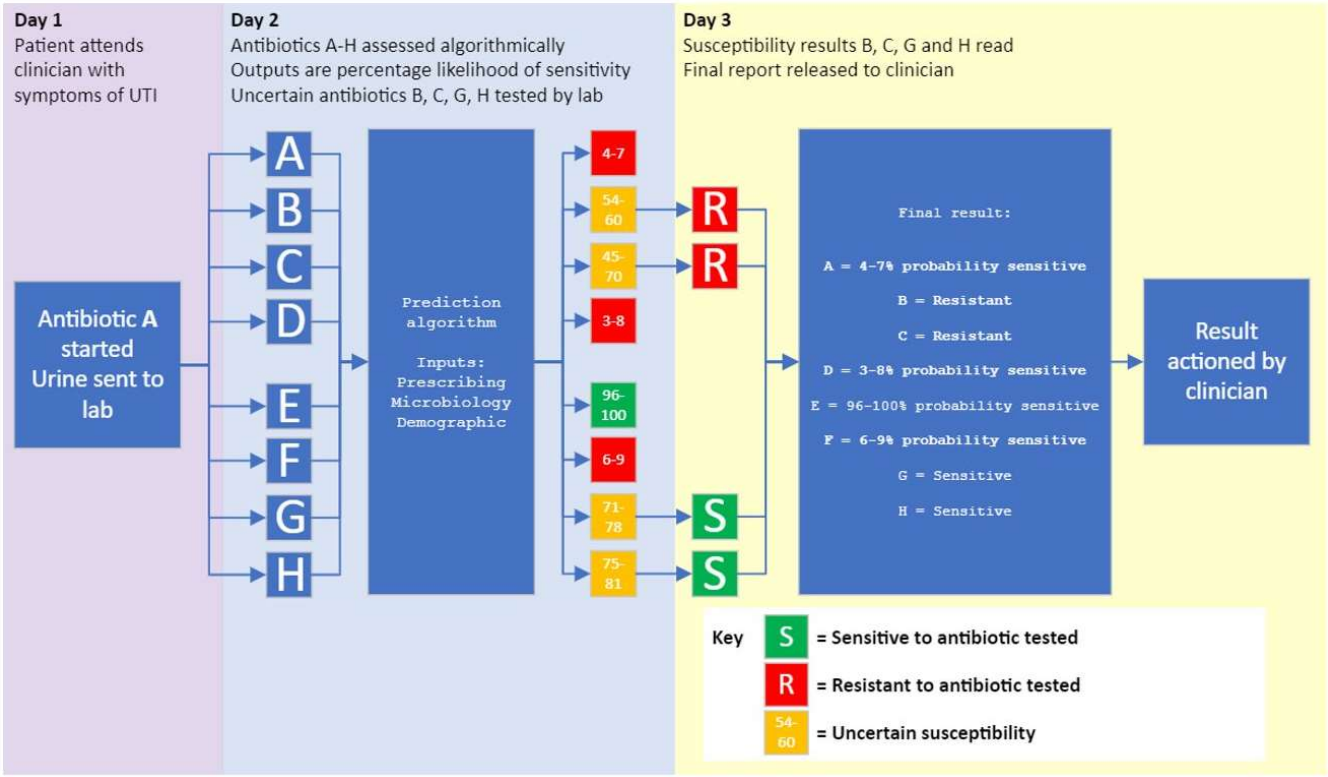

Figure 3: Laboratory workflow and susceptibility testing are reimagined by making use of prior information and prediction algorithms. In this situation, 4 antimicrobials are selected for testing from a panel of 8 for ease of illustration, instead of the panel of 7 used in the study.

The antimicrobials for which susceptibility probability outputs will be created will be those with AST methods available which provide information about the likely activity of antimicrobials which are licensed for the treatment of UTI in the UK. These antimicrobials are listed in Table 2 according to their WHO AWaRe category.

WHO AWaRe Category	**Access**	**Watch**	**Reserve**	**Screening**
Antimicrobials	Amikacin Amoxicillin Amoxicillin-Clavulanate Cefalexin Chloramphenicol Gentamicin Trimethoprim-Sulfamethoxazole Nitrofurantoin	Ceftazidime Ciprofloxacin Ertapenem Erythromycin Fosfomycin (oral) Meropenem Piperacillin-tazobactam Teicoplanin	Aztreonam Ceftazidime-avibactam Ceftolozane-tazobactam	Cefoxitin Cefpodoxime

	Pivmecillinam Trimethoprim	Temocillin Vancomycin		

Table 2: The group of antimicrobials from which ADAPT-AST will select a panel of 7 to test on each urine specimen, stratified by WHO AWaRe category. Cefpodoxime and cefoxitin are also part of the classification, but they are used purely in the laboratory for resistance phenotype screening purposes in Liverpool so have been left out of category as screening antimicrobials here.

A range of different BCI techniques will need to be trialled and compared for this process using open-source R, Stan, and Python coding languages on the open-source R-studio and PyCharm proprietary integrated development software platforms. A range of different BCI techniques will need to be trialled and compared for this process. However, binary LR with Markov Chain Monte Carlo (MCMC) estimation of the posterior distribution is likely to be the most appropriate initial trial technique for this because:

- A predictive model incorporating multiple input variables needs to be developed
- A supervised BCI statistical technique can be used here because the labels of the input variables are known
- The output variable only has two possible outcomes: ‘S’ for susceptible and ‘R’ for resistant
- The output variable can be expressed as a probability of outcome ‘S’ or ‘R’
- Parameters will need to be estimated by random sampling from a posterior probability distribution which is influenced by random variables and in which there is probabilistic dependence between samples(53,54)

Proposed models will be trained on data from a randomly selected 80% of eligible specimens from the retrospective data from between 17th May 2017 and the present. The models will be based on a consistent time frame prior to each urine specimen chosen. This is because firstly the influence of prior data on current antimicrobial susceptibility probability is likely to diminish with increasing time, and secondly it will avoid bias being introduced by more data being available for specimens at the end of the dataset(39). Different periods will be trialled for those which give the highest predictive value. Wherever possible, algorithms will be designed in a ‘modular’ way with intuitive exchangeable components. This will increase the potential scalability and reapplication of algorithms in other settings (e.g., LMICs) with minimal algorithm modification.

3.2.2 Bayesian causal inference Model Testing

The statistical performance of models will then be assessed 'out of sample' on the remaining 20% of specimens in the same retrospective dataset. The techniques used to assess and refine the initial LR-derived model will be developed using R, Stan, and Python programming languages, and will consist of:

- Counterfactual analysis: Input variables will be sequentially removed from and returned to the model to ascertain their likely causative impact, and therefore determine whether they should be included in the final model(27). Where sufficient data are available, sensitivity analyses will also be performed to analyse to ensure that unjustified prediction inequalities do not occur(41).
- Scoring rule loss functions: Different performance assessments will be required depending on the models used. For LR, measures such as log loss score and AUC-ROC will provide measures of model accuracy by quantifying difference between the probability of antimicrobial susceptibility predicted by the model and observed antimicrobial susceptibility(55,56).

3.3 Development and Clinical Evaluation of ADAPT-AST in a simulated clinical setting

3.3.1 AST Decision Making Algorithm Development

The final stage of the project will be evaluating the potential clinical and resource impact of the testing framework on the same 20% of samples used in the testing dataset above. To do this, a clinically suitable AST decision making model first will need to be developed. Comparison of a range of algorithmic approaches is likely to be required, but the aim of the first decision model trialled will be to select 7 antimicrobials for testing (the number in the standard first-line panel currently tested under existing LCL SOPs). The antimicrobials that will be initially eligible for AST selection are those for which probability calculations will be outputted, listed above in Table 2.

The first prioritisation approach trialled will be to rank antimicrobial choices in descending order of susceptibility uncertainty except for cefpodoxime and cefoxitin, then to apply an algorithm in Figure 4 to ensure that a mixture of oral and intravenous agents is tested.

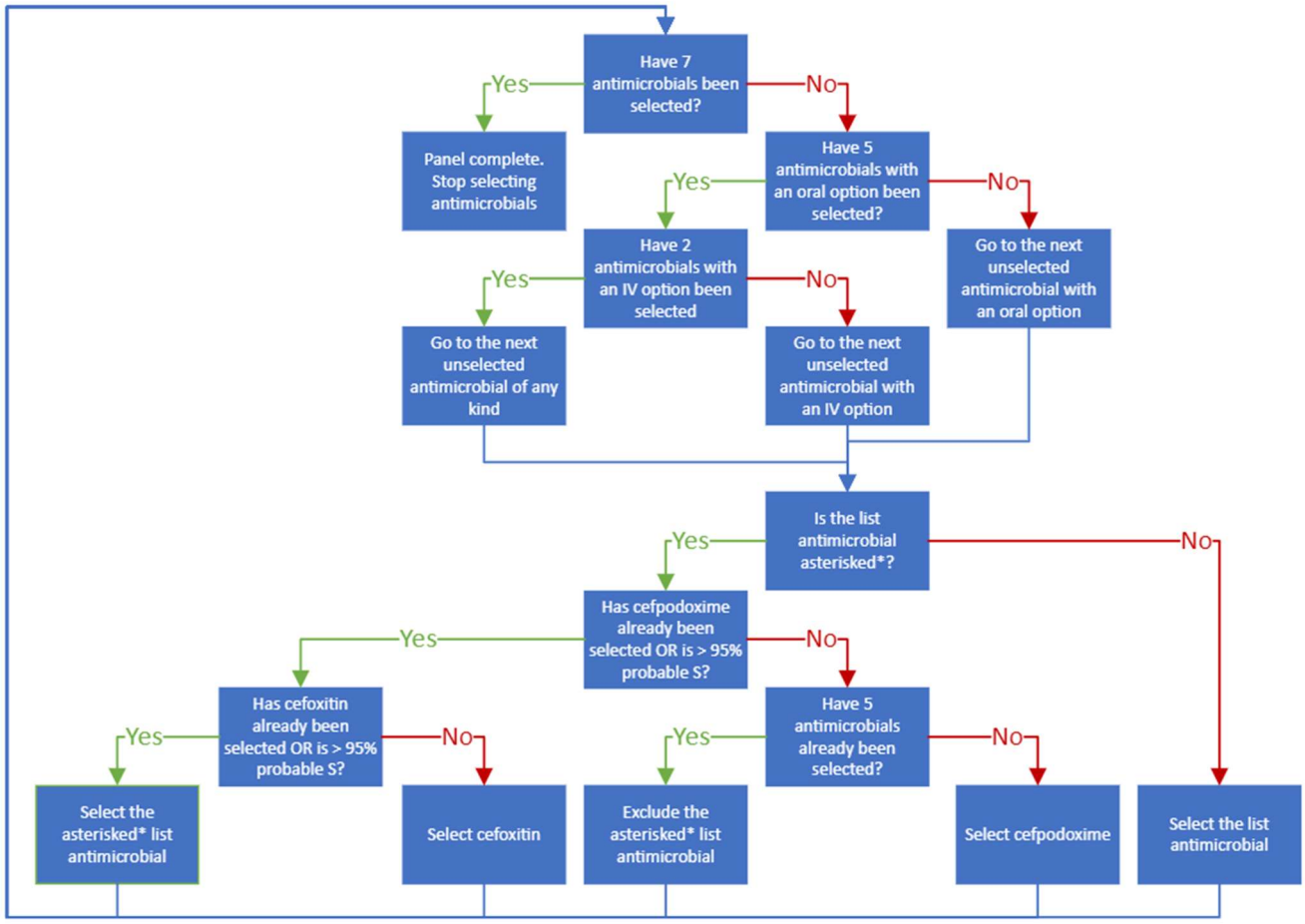

Figure 4: Algorithm used to select 7 antimicrobials from the sorted shortlist.

*Amoxicillin, Amoxicillin-Clavulanate, Aztreonam, Cefalexin, Ceftazidime, Phenoxymethylpenicillin, Piperacillin-tazobactam.

Cefpodoxime or cefoxitin are not used as treatments in the local area, but resistance to either antimicrobial indicates a high probability of resistance mechanisms which preclude use of some beta lactams (BL) and beta-lactam-beta-lactamase-inhibitor combinations (BL-BLIC)(57,58). Therefore, the algorithm will also be designed for cefpodoxime and cefoxitin to only be tested when their susceptibility is sufficiently uncertain (a defined classification threshold will be calculated for this) and there is a chance of the affected BLs / BL-BLICs, cefpodoxime and cefoxitin all being tested as part of the panel of 7.

No additional weighting, ranking or exclusion criteria will initially be applied, but depending on evaluation based on a priori microbiology / clinical / health economic knowledge, measures may need to be taken to further steward the AST choices. These may include exclusion of certain antimicrobial choices which are less practically useful for UTI but for which resistance uncertainty may be high (e.g., teicoplanin) or algorithmic prioritisation/exclusion/weighting measures for factors such as WHO

AWaRe class, nephrotoxicity and risk of *Clostridioides difficile* diarrhoea. Stakeholder consultation by questionnaire will help to inform these sections of the algorithm. Participants in this element of the study will be:

- General Practitioners
- Antimicrobial Pharmacists
- Medical Microbiologists
- Infectious Diseases Physicians
- Acute Medicine Physicians
- Accident and Emergency Physicians
- Bacteriology senior biomedical scientist

Invitations to participate will be disseminated by all-user trust email, with compensation offered for participant time (within sponsor allowances). Participation will be outside time dedicated for clinical care. Formal informed consent of participants will be sought in writing via email, with opportunities for participants to ask questions provided via email and Microsoft Teams. The participant consent form is enclosed in Appendix 1. Once consent has been obtained, participants will be provided with a questionnaire (enclosed in Appendix 2) asking them to rank pseudonymised antimicrobials by order of appropriateness based on several factors. These findings will be analysed by linear regression to output regression coefficients, which will be considered for incorporation into the selection algorithm.

3.3.2 Outcomes

Outcomes will be assessed and compared between the 7-result AST panels provided by ADAPT-AST algorithms and those provided by standard of care algorithms based on LCL current SOPs as of 2023, both of which will be run on every available specimen. Susceptibility information will be assessed alongside drug allergies, renal function, drug interactions to characterise results that 'should', 'could', and cannot be actioned by the clinician. Computer algorithms will be written to perform this assessment in line with the schema in Figure 5 (amoxicillin provided as an example).

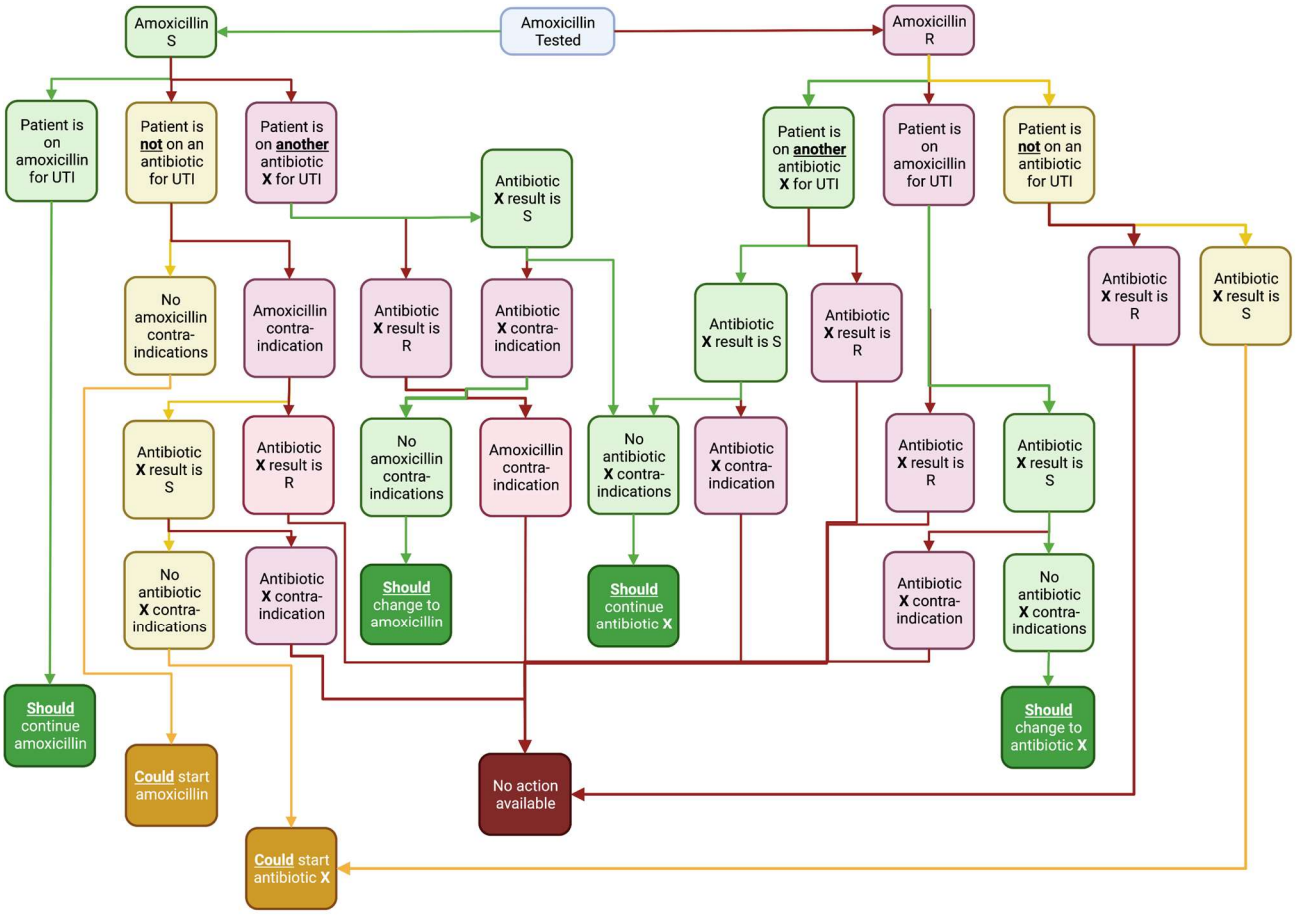

Figure 5: Algorithmic schema for categorisation of antimicrobial results according to whether the clinician 'should', 'could', or cannot action them. In this example, amoxicillin is the result being assessed and antibiotic 'X' refers to any other antibiotic.

Primary outcome:

The number of antimicrobial susceptibility results that 'should' spur action by a clinician on the day that actual first-line AST results were reported (in green on Figure 5).

Secondary outcomes:

5. The number of days until a result that 'should' spur action by a clinician (in green on Figure 5).
6. The number of antimicrobial susceptibility results that 'could' spur action by a clinician on the day that actual first-line AST results were reported (in yellow on Figure 5).
7. The number of days until a result that 'could' spur action by a clinician (in yellow on Figure 5).

8. The projected health economic cost per specimen, including laboratory (e.g., consumable cost) and patient (e.g., drug toxicity, clinical failure) measures guided by the above criteria.

3.3.3 Statistical Analysis

A chi square goodness of fit test will be conducted to determine whether the counts of, and number of days until, 'should' and 'could' action results follow a Poisson distribution; if so, a one-sample exact Poisson test will be run to determine whether the mean number of 'should' and 'could' action results per specimen and the number of days to these results using ADAPT-AST would be different from the expected number using the LCL SOP. The mean financial cost per specimen will be analysed either with a single-sample paired t-test or Wilcoxon signed-ranks test depending on whether the data are normally distributed.

3.3.4 Potential Limitations of Simulated Interventional Study

Given that the intervention in this study is simulated, there is no risk of harm to patients through incorrect testing decisions. However, the lack of intervention means that the impact on the eventual prescription cannot be assessed, because prescriber behaviour is not incorporated. Future randomised interventional studies will therefore need to be performed to assess impact before the approach could be deployed in the clinical setting. Another potential limitation is missing data, either from patients who have been transferred from outside of the region, or who have passed through areas without electronic prescribing data (e.g., Accident and Emergency, Intensive Care). Although this will be a small proportion of the antibiotics prescribed and administered, the potential impact of this will need to be considered when the model is explained. An ability to distinguish between transfer from out of area (for example, a high AMR setting) and patients presenting to healthcare for the first time may also affect the predictive value of an apparent lack of healthcare exposure for a lower AMR probability by diluting its effect. Some antimicrobial susceptibility data will also be missing due to agents not having been tested; the impact of this will need to be weighted within the process of algorithm development.

3.4 Economic evaluation of ADAPT-AST

A decision model will be developed to estimate the cost and health outcomes associated with using ADAPT-AST to inform clinical decisions about the treatment of UTI in hospitals, compared to standard care (as observed within the Liverpool data). The introduction of ADAPT-AST is intended to allow patients to receive earlier access to appropriate therapy by better targeting the use of antibiotic susceptibility tests. This has the potential to identify early a larger set of antibiotics to which an individual is susceptible. This avoids the cost and health impacts of delayed appropriate therapy, and by expanding the choice of therapies allows clinical teams to prioritise between these therapies based on other important factors: namely toxicity profile; mode of administration; risk of AMR; and anticipated differences in efficacy.

The economic evaluation will compare the AST Decision Making Algorithms evaluated within the simulated interventional trial to standard care. A conceptual model will be developed in consultation with relevant stakeholders (e.g. clinicians who treat UTIs, microbiologists) to identify how the AST information available under different AST Decision Making Algorithms (and under standard care) is likely to influence antibiotic prescription, and how this is likely to modify costs and outcomes amongst patients treated for UTIs in UK secondary care. The conceptual model will also identify which health and cost effects of introducing ADAPT-AST should be prioritised for modelling based on their likely importance for patient health and costs. This conceptual model will then form the basis for a more detailed model protocol documenting the model structure and data sources for the quantitative model.

The susceptibility information expected to be available to inform clinical decision making under ADAPT-AST and standard care will be obtained from the clinical prediction model and simulated interventional study. An important consideration for the economic evaluation will be how this information influences clinical decision making. This will be informed by data from the Liverpool area dataset where feasible, supplemented by the literature and clinical opinion where necessary. The decision model will reflect the short-term costs of introducing ADAPT-AST in terms of clinical time, infrastructure and microbiology tests. The long-term cost implications of using ADAPT-AST to guide clinical decision making will be reflected accounting for impacts on length of stay and intensity of care within hospitals (e.g. requirement for IV antibiotics, significant treatments for antibiotic-related toxicity, use of ICU/HDU). The short- and long-term implications of using ADAPT-AST for patient health will be quantified as quality-adjusted life years (QALYs) which account for both morbidity and mortality effects, and can reflect both efficacy and toxicity effects of antibiotics. We will separately quantify the expected impact of using ADAPT-AST and standard care on the proportion of patients receiving treatment with an antibiotic falling under the WHO Access, Watch and Reserve categories. Categorisation of patients receiving multiple agents will be established in discussion with experts. Cost, quality of life and mortality parameters will be obtained from a combination of the Liverpool area dataset, the literature and expert opinion where necessary.

Model results will be summarised as incremental cost-effectiveness ratios (ICERs) and net health effects and an assessment of whether ADAPT-AST represents a cost-effective use of NHS resources will be provided, appropriately accounting for health opportunity costs. The sensitivity of the model results to important data inputs and assumptions will be tested using deterministic sensitivity analysis and probabilistic sensitivity analysis.

4 ETHICAL AND REGULATORY CONSIDERATIONS

4.1 Assessment and management of risk

There is no interventional component to the study; there is therefore no risk of direct harm to individuals. The risks to consider at study outset are breaches in Information Governance, Data Protection and Data Security. The processes for obtaining appropriate regulatory approval and ensuring that such breaches do not occur are detailed in sections 4.2.1 and 4.6 of this protocol. Six-monthly central monitoring milestones of risk will be put in place at which point the risk assessment will be updated. Risk assessments will also be updated if a protocol deviation occurs. If data quality or quantity is insufficient at the first annual milestone, a substantial amendment will be submitted to amend the route of data access.

4.2 Regulatory Review and Amendments

4.2.1 Ethical and Regulatory Review, Compliance and Reports

The study will be conducted in accordance with the principles of Good Clinical Practice (GCP). Before the start of the study, LUHFT Sponsorship, Cheshire and Merseyside ICB/CIPHA data asset and access group (DAAG – see Section and research ethics committee (REC) approval will be secured by the CI.

4.2.2 Amendments

A study management group (SMG) will be convened and chaired by the CI with representation from across the study team and a public lay representative to provide ongoing management and oversight, support amendments and deal with reporting, data issues and protocol deviations. It is the sponsor's responsibility to decide whether an amendment is substantial or non-substantial. If a substantial amendment to the protocol is required, the chief investigator (CI) or delegate will submit a valid notice of amendment to the sponsor for consideration via the appropriate route. Amendment history will be tracked by keeping all protocol versions in a secure password-protected folder in the LUHFT computer system. Substantial amendments will not be implemented until all necessary regulatory approvals have been formally approved.

4.3 Peer review

The study design was peer reviewed by Dr George Drusano, Director of the Institute for Therapeutic Innovation at the University of Florida, and Dr Ang Li, Consultant in Medical Microbiology at Liverpool University Hospitals NHS Foundation Trust. The reviewer(s) are not involved in the study in any way, and have sufficient knowledge of the clinical subject area to consider the protocol's clinical, methodological and service aspects.

4.4 Patient & Public Involvement

In 2022, an AMR Citizens' Jury was commissioned by the University of Liverpool to explore public and patient perspectives on collaborative monitoring of AMR data for applications of this kind in the Liverpool area(36). Jurors were recruited by advertisement and selected to provide a representative cohort of the Liverpool city area. They were provided with subject matter background on AMR, data protection and information governance by domain experts over several days. They were then asked to discuss and provide their assessment of the acceptability, design, management, undertaking, research, analysis and dissemination of the results of integrated data systems used to facilitate clinical care, drug development and research. The jury was broadly supportive of the use of integrated regional pseudonymised data for this process and raised important points about the legal, security and quality frameworks that will be required to underpin healthcare data projects of this kind.

4.5 Protocol compliance

All Protocol deviations will be recorded and reviewed as part of regular SMG meetings and reported to Sponsor where deemed appropriate.

4.6 Consent, data protection and patient confidentiality

Pre-existing Population Health DPIAs have been signed by data controllers for healthcare data to be transferred to the ICB/CIPHA secure data environment (SDE) within Cheshire and Merseyside ICB firewalls, to perform routine data linkage and pseudonymisation. Only pseudonymised, non-identifiable data will be accessed by researchers with approval from ICB/CIPHA DAAG. Patients who opted out of

data sharing for purposes other than direct care (Type 1 objections) are excluded from the flow of data into the ICB/CIPHA SDE. The study will adhere to LUHFT and CIPHA governance structures for data protection and patient confidentiality to ensure compliance to legal standards. As study sponsor, Liverpool University Hospitals NHS Foundation Trust is the overriding governance structure.

4.6.1 Data Access via CIPHA

The data access and cleaning processes are detailed in Section 3.1. Core project datasets will be curated and maintained from the integrated healthcare records detailed in section 3.1.2, where available. An application for access to the data within CIPHA will be made through the CIPHA Data Asset and Access Group (DAAG). CIPHA are an NHS Population Health platform established during the COVID-19 pandemic which will provide the software, tools, compute, and governance for project access. Approval will be obtained from DAAG for CIPHA to permit access of the ADAPT-AST team to pseudonymised data (pseudonymised by NHS Data Services for Commission Regional Offices [DSCRO] according to standard policies) in the CIPHA secure data environment. This environment is administered by CIPHA. CIPHA meets requirements of GDPR, The Data Protection Act 2018 and the NHS Data Security and Protection Toolkit, and is certified to ISO27001, ISO9001 and Cyber Essentials standards.

The legal basis for the study under General Data Protection Regulation (GDPR) is Article 9(2)(j), in that data processing and access is necessary for the purpose of scientific or statistical purposes in accordance with Article 89(1). Data controllers, the third party (NHS Cheshire and Merseyside ICB/CIPHA) and investigators must comply with the requirements of the Data Protection Act 1998 with regards to the collection, storage, processing, and disclosure of personal information and will uphold the Act's core principles. Only the minimum amount of data pertaining to the minimum number of individuals required to facilitate the study methodology will be used. Pseudonymisation keys are kept secure within an encrypted, password protected ICB/CIPHA data warehouse and only accessible by an ICB/CIPHA engineer assigned to data linkage and pseudonymisation (not accessible by the research team). The details of the pseudonymisation process are listed in Section 3.1.4. The data flow diagram for ADAPT-AST is displayed in Figure 6.

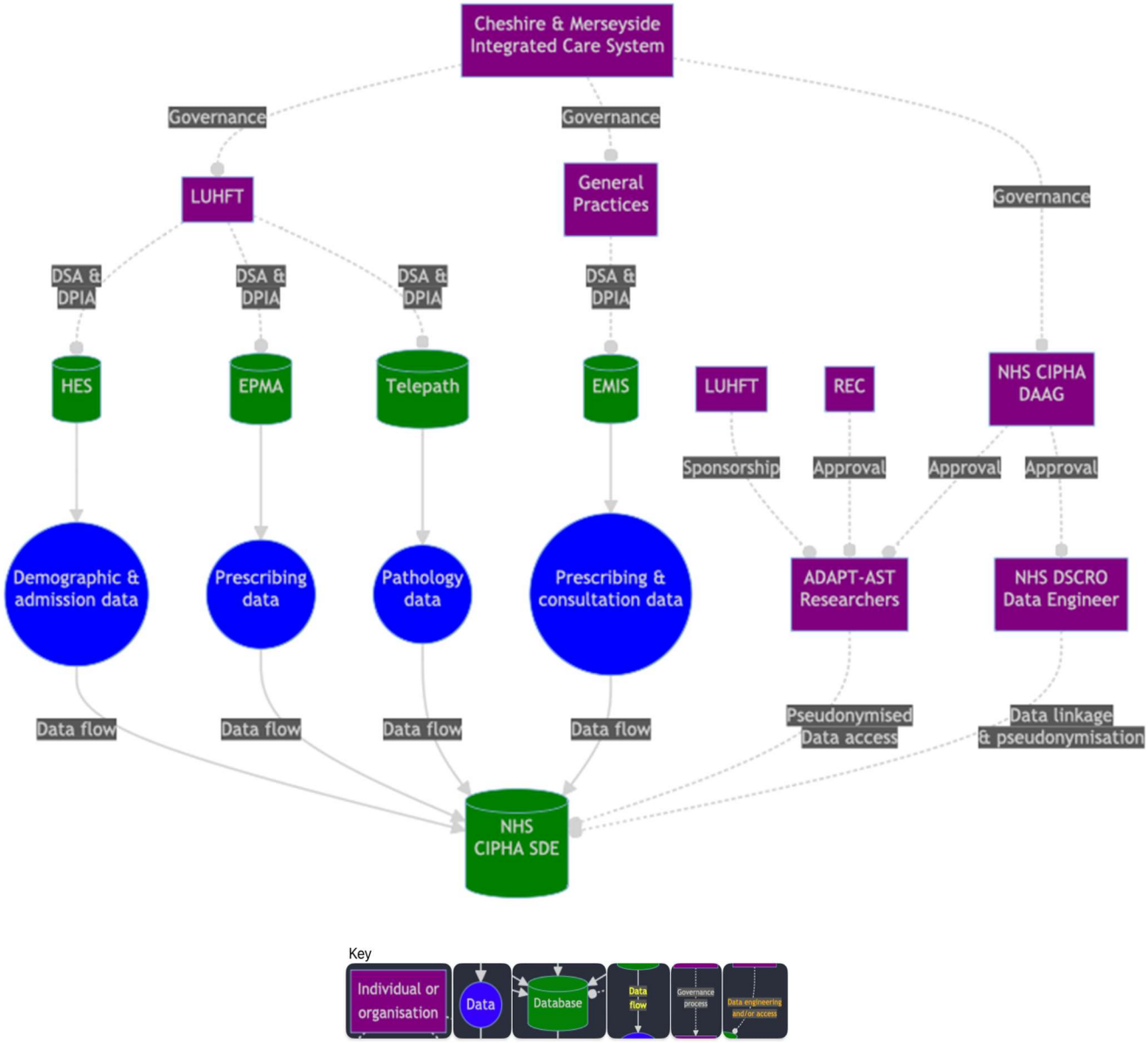

Figure 6: ADAPT-AST data flow diagram. DAAG and LUHFT approvals will facilitate portal access to pseudonymised LUHFT and GP data within the NHS CIPHA SDE.

Access to the pseudonymised datasets will be limited to members and collaborators of the study group. Access will be password protected by NHS or UoL IT security. The pseudonymised data will be stored for 10 years after publication, in accordance with the UK Research and Innovation (UKRI) best practice recommendations (UK Research and Innovation, 2018). A data catalogue will be maintained. The data custodian will be the CI, and will be responsible for the safe custody, transport and storage of any aggregate data resulting from this project.

4.6.2 Confidentiality

The Chief Investigator will preserve the confidentiality of participants taking part in the study and will abide by the Data Protection Act 2018 and the UK GDPR as amended from time to time and any successor legislation in the UK and any other directly applicable regulation relating to data protection and privacy.

4.6.3 Audits

The study may be subject to inspection and audit by LUHFT under their remit as sponsor and other regulatory bodies to ensure adherence to GCP and the UK Policy Framework for Health and Social Care Research (v3.2 10th October 2017).

4.7 Indemnity

LUHFT holds Indemnity and insurance cover which will apply to this study.

4.8 Access to the final study dataset

The final study dataset will be accessible to the CI and the data scientists performing the data analysis, and the CIPHA engineer(s) facilitating access to the pseudonymised data.

5 END OF STUDY

The end of study will be when analysis of the simulated interventional study data on the final ADAPT-AST algorithm version is complete, at which point an End of Study Declaration will be completed and submitted to the Sponsor. After the end of study is declared no study activity, other than final analysis of the data (following 'lock' of the study database) and report writing, will be undertaken.

6 DISSEMINATION POLICY

6.1 Dissemination policy

The data arising from the study will be owned by the study authors. On completion of the study, the data will be analysed and tabulated, and a Final Study Report prepared, which will be accessible on the Centres for Antimicrobial Optimisation Network (CAMO-NET) website. Participating investigators

can publish any of the study data with permission of the CI. Public participants of the Liverpool AMR Citizens' Jury will be informed of the outcome of the study by provision of the publication. All statistical and BCI code used to generate the results will be made available open source, shared via a GitHub public repository following journal publication.

6.2 Authorship eligibility guidelines and any intended use of professional writers

The main study results will be published as soon as a manuscript is completed, in the name of the study in a peer-reviewed journal, on behalf of all collaborators. The manuscript will be prepared by a writing group composed of the investigators. All investigators will be granted authorship on the final study report in line with International Committee for Medical Journal Editors criteria for individually named authors and group authorship.

7 ARCHIVING

Data and all appropriate documentation will be stored for a minimum of 10 years after the completion of the study within the password-encrypted CIPHA secure data environment, including the follow-up period.

8 REFERENCES

1. Smith DRM, Pouwels KB, Hopkins S, Naylor NR, Smieszek T, Robotham JV. Epidemiology and health-economic burden of urinary-catheter-associated infection in English NHS hospitals: a probabilistic modelling study. Journal of Hospital Infection. 2019 Sep;103(1):44–54.
2. Medina M, Castillo-Pino E. An introduction to the epidemiology and burden of urinary tract infections. Ther Adv Urol. 2019 Jan 2;11:175628721983217.
3. Paul R. State of the globe: Rising antimicrobial resistance of pathogens in urinary tract infection. J Glob Infect Dis. 2018;10(3):117.
4. Pokharel S, Raut S, Adhikari B. Tackling antimicrobial resistance in low-income and middle-income countries. BMJ Glob Health. 2019 Nov 10;4(6):e002104.
5. MAST. MAST URI®SYSTEM [Internet]. [cited 2022 Dec 21]. Available from: https://mast-group.com/uk/products/mast-uri-system/
6. ThermoFisher. Sensititre™ Gram Negative MIC Plate [Internet]. [cited 2022 Dec 21]. Available from: https://www.thermofisher.com/order/catalog/product/GN2F
7. Opatowski M, Brun-Buisson C, Touat M, Salomon J, Guillemot D, Tuppin P, et al. Antibiotic prescriptions and risk factors for antimicrobial resistance in patients hospitalized with urinary

tract infection: a matched case-control study using the French health insurance database (SNDS). BMC Infect Dis. 2021 Dec 14;21(1):571.

8. Galar A, Leiva J, Espinosa M, Guillén-Grima F, Hernáez S, Yuste JR. Clinical and economic evaluation of the impact of rapid microbiological diagnostic testing. Journal of Infection. 2012 Oct;65(4):302–9.

9. Zhu H, Chen Y, Hang Y, Luo H, Fang X, Xiao Y, et al. Impact of inappropriate empirical antibiotic treatment on clinical outcomes of urinary tract infections caused by Escherichia coli: a retrospective cohort study. J Glob Antimicrob Resist. 2021 Sep;26:148–53.

10. Dadgostar P. Antimicrobial Resistance: Implications and Costs. Infect Drug Resist. 2019 Dec;Volume 12:3903–10.

11. Barlow G, Nathwani D, Davey P. The CURB65 pneumonia severity score outperforms generic sepsis and early warning scores in predicting mortality in community-acquired pneumonia. Thorax. 2007 Mar 1;62(3):253–9.

12. Klastersky J, Paesmans M, Rubenstein EB, Boyer M, Elting L, Feld R, et al. The Multinational Association for Supportive Care in Cancer Risk Index: A Multinational Scoring System for Identifying Low-Risk Febrile Neutropenic Cancer Patients. Journal of Clinical Oncology. 2000 Aug 16;18(16):3038–51.

13. Selby J v., Fireman BH. Building Predictive Models for Clinical Care—Where to Build and What to Predict? JAMA Netw Open. 2021 Jan 22;4(1):e2032539.

14. Chen L. Overview of clinical prediction models. Ann Transl Med. 2020 Feb;8(4):71.

15. Wood A, Denholm R, Hollings S, Cooper J, Ip S, Walker V, et al. Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. BMJ. 2021 Apr 7;n826.

16. Yamin M. Information technologies of 21st century and their impact on the society. International Journal of Information Technology. 2019 Dec 16;11(4):759–66.

17. Shamout F, Zhu T, Clifton DA. Machine Learning for Clinical Outcome Prediction. IEEE Rev Biomed Eng. 2021;14:116–26.

18. van Calster B, Verbakel JY, Christodoulou E, Steyerberg EW, Collins GS. Statistics versus machine learning: definitions are interesting (but understanding, methodology, and reporting are more important). J Clin Epidemiol. 2019 Dec;116:137–8.

19. Fitzpatrick F, Tarrant C, Hamilton V, Kiernan FM, Jenkins D, Krockow EM. Sepsis and antimicrobial stewardship: two sides of the same coin. BMJ Qual Saf. 2019 Sep;28(9):758–61.

20. Mancini A, Vito L, Marcelli E, Piangerelli M, de Leone R, Pucciarelli S, et al. Machine learning models predicting multidrug resistant urinary tract infections using “DsaaS”. BMC Bioinformatics. 2020 Aug 21;21(Suppl 10):347.

21. Kim C, Park RW, Rhie SJ. 174. Development of a Machine Learning Prediction Model to Select Empirical Antibiotics in Patients with Clinically Suspected Urinary Tract Infection using Urine Culture Data. Open Forum Infect Dis. 2021 Dec 4;8(Supplement_1):S196–S196.

22. 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 Nov 4;12(568).

23. Tandan M, Timilsina M, Cormican M, Vellinga A. Role of patient descriptors in predicting antimicrobial resistance in urinary tract infections using a decision tree approach: A retrospective cohort study. Int J Med Inform. 2019 Jul 1;127:127–33.

24. 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 Jul 31;

25. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019 Dec 29;17(1):195.

26. Lueangarun S, Leelarasamee A. Impact of inappropriate empiric antimicrobial therapy on mortality of septic patients with bacteremia: a retrospective study. Interdiscip Perspect Infect Dis. 2012;2012:765205.

27. Prosperi M, Guo Y, Sperrin M, Koopman JS, Min JS, He X, et al. Causal inference and counterfactual prediction in machine learning for actionable healthcare. Nat Mach Intell. 2020 Jul 13;2(7):369–75.

28. Burton RJ, Albur M, Eberl M, Cuff SM. Using artificial intelligence to reduce diagnostic workload without compromising detection of urinary tract infections. BMC Med Inform Decis Mak. 2019 Dec 23;19(1):171.

29. Peiffer-Smadja N, Dellière S, Rodriguez C, Birgand G, Lescure FX, Fourati S, et al. Machine learning in the clinical microbiology laboratory: has the time come for routine practice? Clinical Microbiology and Infection. 2020 Oct;26(10):1300–9.

30. Begoli E, Bhattacharya T, Kusnezov D. The need for uncertainty quantification in machine-assisted medical decision making. Nat Mach Intell. 2019 Jan 7;1(1):20–3.

31. Lord Carter of Coles. Operational productivity and performance in English NHS acute hospitals: Unwarranted variations – An independent report for the Department of Health by Lord Carter of Coles – February 2016 [Internet]. 2016 [cited 2022 Oct 28]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/499229/Operational_productivity_A.pdf

32. Sharland M, Cappello B, Ombajo LA, Bazira J, Chitatanga R, Chuki P, et al. The WHO AWaRe Antibiotic Book: providing guidance on optimal use and informing policy. Lancet Infect Dis. 2022 Nov;22(11):1528–30.

33. NHS England. NHS Standard Contract 2022/23 Technical Guidance [Internet]. 2022 [cited 2022 Oct 28]. Available from: https://www.england.nhs.uk/wp-content/uploads/2022/03/nhs-standard-contract-22-23-technical-guidance-march-25-update.pdf

34. NHS England. Commissioning for Quality and Innovation (CQUIN): 2022/23 Guidance [Internet]. 2022 [cited 2022 Oct 28]. Available from: https://www.england.nhs.uk/wp-content/uploads/2022/01/B1477-i-cquin-22-23-march-2022.pdf

35. HM Government. Tackling antimicrobial resistance 2019–2024: The UK's five-year national action plan. 2019.

36. Center for Democratic Processes. Liverpool AMR Citizens' Jury Full Report Liverpool Antimicrobial Resistance (AMR) Citizens' Jury [Internet]. 2022 [cited 2022 Oct 28]. Available from: https://www.liverpool.ac.uk/media/livacuk/iib/AMR_Full,Report_Final,V1.4.pdf

37. Allel K, García P, Labarca J, Munita JM, Rendic M, Grupo Colaborativo de Resistencia Bacteriana, et al. Socioeconomic factors associated with antimicrobial resistance of Pseudomonas aeruginosa, Staphylococcus aureus, and Escherichia coli in Chilean hospitals (2008–2017). Revista Panamericana de Salud Pública. 2020 Sep 23;44:1.

38. Struelens MJ. The epidemiology of antimicrobial resistance in hospital acquired infections: problems and possible solutions. BMJ. 1998 Sep 5;317(7159):652–4.

39. Costelloe C, Metcalfe C, Lovering A, Mant D, Hay AD. Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis. BMJ. 2010 Jun 2;340(may18 2):c2096–c2096.

40. Bonten MJM, Weinstein RA. The Role of Colonization in the Pathogenesis of Nosocomial Infections. Infect Control Hosp Epidemiol. 1996 Mar;17(3):193–200.

41. Global Future Council on Human Rights 2016-2018. How to Prevent Discriminatory Outcomes in Machine Learning. Geneva; 2018 Mar. Report No.: REF 120318-case 00040065.

42. HM UK Government. Discrimination: your rights: Types of Discrimination ('protected characteristics') [Internet]. [cited 2023 Jan 23]. Available from: https://www.gov.uk/discrimination-your-rights

43. Bujang MA, Sa'at N, Tg Abu Bakar Sidik TMI, Chien Joo L. Sample Size Guidelines for Logistic Regression from Observational Studies with Large Population: Emphasis on the Accuracy Between Statistics and Parameters Based on Real Life Clinical Data. Malaysian Journal of Medical Sciences. 2018;25(4):122–30.

44. Lu J, Liu A, Dong F, Gu F, Gama J, Zhang G. Learning under Concept Drift: A Review. IEEE Trans Knowl Data Eng. 2018;1–1.

45. Sampson R, Cooper J, Barbour R, Polson R, Wilson P. Patients' perspectives on the medical primary–secondary care interface: systematic review and synthesis of qualitative research. BMJ Open. 2015 Oct 15;5(10):e008708.

46. Combined Intelligence for Population Health Action. CIPHA [Internet]. [cited 2023 Jan 23]. Available from: https://www.cipha.nhs.uk

47. Civic Data Cooperative. Liverpool City Region Civic Data Cooperative [Internet]. [cited 2023 Jan 23]. Available from: https://civicdatacooperative.com

48. Microsoft. Azure Data Factory Documentation [Internet]. [cited 2023 Jan 23]. Available from: https://learn.microsoft.com/en-us/azure/data-factory/

49. Green MA, Hungerford DJ, Hughes DM, Garcia-Fiñana M, Turtle L, Cheyne C, et al. Changing patterns of SARS-CoV-2 infection through Delta and Omicron waves by vaccination status,

previous infection and neighbourhood deprivation: a cohort analysis of 2.7 M people. BMC Infect Dis. 2022 Nov 26;22(1):889.

50. European Committee on Antimicrobial Susceptibility Testing. EUCAST Expected Resistant Phenotypes [Internet]. 2022 [cited 2022 Dec 14]. Available from: https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Expert_Rules/2022/Expected_Resistant_Phenotypes_v1.1_20220325.pdf

51. European Committee on Antimicrobial Susceptibility Testing. EUCAST Redefining susceptibility testing categories S, I and R [Internet]. 2019 [cited 2022 Dec 15]. Available from: https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/EUCAST_Presentations/2018/EUCAST_-_Intermediate_category_-_information_for_all.pdf

52. Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Comput Sci. 2021 May 22;2(3):160.

53. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. J Biomed Inform. 2002 Oct;35(5–6):352–9.

54. van Ravenzwaaij D, Cassey P, Brown SD. A simple introduction to Markov Chain Monte–Carlo sampling. Psychon Bull Rev. 2018 Feb 11;25(1):143–54.

55. Vasandani J, Bharti S, Singh D, Priyadarshi S. Improving Model Accuracy with Probability Scoring Machine Learning Models. In 2021. p. 517–30.

56. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit. 1997 Jul;30(7):1145–59.

57. Harris PNA, Tambyah PA, Lye DC, Mo Y, Lee TH, Yilmaz M, et al. Effect of Piperacillin-Tazobactam vs Meropenem on 30-Day Mortality for Patients With *E coli* or *Klebsiella pneumoniae* Bloodstream Infection and Ceftriaxone Resistance. JAMA. 2018 Sep 11;320(10):984.

58. European Committee on Antimicrobial Susceptibility Testing. EUCAST guidelines for detection of resistance mechanisms and specific resistances of clinical and/or epidemiological importance [Internet]. 2013 [cited 2022 Dec 16]. Available from: https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Resistance_mechanisms/EUCAST_detection_of_resistance_mechanisms_v1.0_20131211.pdf

9 APPENDICES

9.1 Health Professional Questionnaire Consent Form

Title of Study: Adaptive Prediction of Antimicrobial Susceptibility and its Implementation to Improve the Management of Urinary Tract Infection

Principal Investigator: Alex Howard

Thank you for responding to the invitation to participate in ADAPT-AST as an expert healthcare professional. This questionnaire aims to explore your perception of the importance of various attributes of antimicrobials used in the management of urinary tract infection. Before you decide whether to participate, it is important for you to understand the purpose, procedures, and risks of the study. Please read this form carefully and ask any questions you may have before agreeing to participate.

Purpose of the study:
The aim of this study is to develop and evaluate an adaptive informatics approach for laboratory antimicrobial susceptibility testing (AST) for urinary tract infection (UTI) pathogens compared with current practice, as a potential means to improve patient outcomes, reduce AMR risks and reduce waste of laboratory resources.

Procedures:
If you agree to participate in this study, you will be asked to complete a questionnaire which will take up an estimated 15 to 30 minutes of your time This will be stored alongside a general description of your role (no personal identifying information will be stored).

Risks:
There are minimal risks associated with participating in this study. We will not specifically ask questions related to how you personally have managed individual patients. Please inform the researcher if you have any questions or concerns before participating in the study.

Benefits:
This study will give you the opportunity to contribute to medical research and the possibility of gaining a deeper understanding of perceptions related to antimicrobial treatment of UTI.

Confidentiality:
Responses will be confidential and pseudonymised. Only general area of professional expertise (e.g., Medical Microbiologist) will be disclosed in study findings.

Voluntary participation:
Participation in this study is completely voluntary. You may refuse to participate or withdraw your participation at any time without penalty or loss of benefits to which you are otherwise entitled. Study participation must be outside of time that is rostered for delivery of clinical care.

Contact information:

If you have any questions or concerns about the study, you may contact the chief investigator alex.howard@liverpoolft.nhs.uk

Consent:
By signing below, you indicate that you have read and understood the information provided in this consent form and agree to participate in this study.

Participant Signature

Date

9.2 Appendix 2 – Health Professional Questionnaire

1. A patient presents to General Practice with clinical features of urinary tract infection. Please rank the following fictional antibiotic treatments from most to least preferred in this scenario by dragging and dropping boxes or using the drop down menus on the left.

	AWaRe class*	*C difficile* risk	Toxicity risk	UTI specificity	Oral option	IV option	Cost
Decloxone	Watch	Low	High	No	No	Yes	High
Choriotroban	Access	Low	High	No	No	Yes	Low
Acetemran	Watch	Low	Low	Yes	Yes	No	Low
Olanzasys	Watch	High	Low	No	No	Yes	Low
Pansolid	Access	Low	Low	No	Yes	No	Low
Protestryl	Watch	High	Low	No	Yes	Yes	Low
Endoxolol	Access	Low	Low	Yes	Yes	No	Low
Adenomadin	Access	High	Low	No	Yes	Yes	Low
Abelfenide	Reserve	Low	Low	No	Yes	Yes	High
Dexaset	Access	Low	Low	No	Yes	Yes	Low
Adrevenac	Watch	High	Low	No	No	Yes	High
Cormide	Watch	Low	Low	No	No	Yes	High
Amrodine	Reserve	High	Low	No	No	Yes	High

9.3 Appendix 3 - Data Management plan

What stage are you at in this project?

- Post-award

Which faculty do you belong to?

- Health and Life Sciences

Do you have, or will you be applying for Ethics approval for your project?

- Yes

Will be you collecting and storing personal or sensitive data as defined under the terms of GDPR? (this includes email addresses, phone numbers, etc)

- No

Will you require space on the Active DataStore?

- No

If you are not using the ADS, where will you store your data?

- Cloud Store

Will you be depositing your data in an open repository at the conclusion of your project?
- No

Your research data

What types of data will be collected or created?
The project will analyse linked, pseudonymised demographic, past medical, hospital admission, general practice consultation, prescribing, microbiology and other laboratory pathology specimen data.

What formats will you use?
Comma Separated Values files

How much data do you estimate you will be collecting and storing?
~ 80-100 columns
~ 250,000-1,000,000 rows
Estimated file size would be between 3GB and 4GB

Documentation

Are there any standards for organising, labelling or describing research data in your field of research. If so, detail below.
Systematised Nomenclature of Medicine Clinical Terms (SNOMED-CT), an international clinical terminology metadata library.

Ethics and Intellectual Property

Who owns the data you will be using, creating or collecting?
Data will be owned by Cheshire and Merseyside Integrated Care Board who manage the CIPHA service and access to this dataset. Prior data protection impact assessments and data sharing agreements have been completed by University Hospitals NHS Foundation Trust and all participating general practices.

Are there any legal, ethical or commercial considerations? If so, how do you propose to deal with them?
The possible risk of reidentification is negligible owing to the pseudonymisation process – project team members must take personal responsibility for data security, working in a way that prevents legal and procedural breaches. Project team members will ensure that statistical disclosure controls are in place on output data to ensure that reidentification is not possible from results.

Storage and Organisation

Where will the data be stored during your project? If you are not using UoL managed drives, explain why.
The NHS Cheshire and Merseyside Integrated Care board / Combined Intelligence for Public Health Action (CIPHA) secure secure data environment.

Are there any security issues relating to the storage of the data?
The data will be de-identified and will not leave the NHS-managed and provided secure environment. Only named and approved ADAPT-AST prroject members will be given access to the data.

Who else will have access to this data during the project?
NHS ICB-employed data engineers within the NHS ICB/CIPHA secure data environment.

Data Sharing

Will you be able to share any of your data in an open access repository?
Statistically derived aggregate data may be shared in publication.

If not 'open', who could have access to your data and how would this be facilitated?
Any shareable data could be shared using the Liverpool Data Catalogue, with any share data directed to by a data statement in any publications.

What formats do you anticipate the data will be shared in?
Not applicable

Long Term Archiving

Will there be any data that you cannot share but will need to be retained in the long term?
Pseudonymised study data and code key will be archived within a password-encrypted secure data environment within the ICB/CIPHA SDE for 10 years.

Where will the data be archived at the end of the project and how long will it be retained?
The research team will request that data be archived for 10 years within the NHS CIPHA Secure Data Environment.

What formats do you anticipate the data will be archived in?
Not applicable - formatting will be set by CIPHA