

Use data linkage to support AMR research through a whole-systems approach

Nina Zhu
CAMO-Net Data Theme Meeting, Kampala, Uganda
06 June 2024

Overview

- Welcome
- Data linkage: why, how, what to achieve within CAMO-Net
- Ongoing work in CAMO-Net UK
- Activity

About me

- Biomedical engineering
- Public health
- Epidemiology
- Health economics



English surveillance programme for antimicrobial utilisation and resistance (ESPAUR)

Report 2020 to 2021





Overview

- Welcome
- Data linkage: why, how, what to achieve within CAMO-Net
- Ongoing work in CAMO-Net UK
- Activity

Imperial College London Proportion of countries reporting: 📕 Large decrease 🧾 Moderate decrease 🔝 No impact 🗾 Moderate increase 📘 Large increased. Reported impact of COVID-19 on laboratory supplies and equipment for AMR activities a. Reported impact of COVID-19 on funding for AMR activities Availability of quality laboratory reagents/consumables for bacteriology and 58% 41% antimicrobial susceptibility testing (n=71) Availability of funding for AMR surveillance at 41% Ability of laboratories to service their machines the national level (n=70) and equipment, e.g. repairs, compliance and 52% 46% updates (n=67) Availability of funding for AMR surveillance at 43% the local (facility) level (n=69) Access to advanced technologies e.g. molecular 50% testing for multidrug resistant organisms (n=66) Why b. Reported impact of COVID-19 on partnerships and oversight for AMR activities e. Reported impact of COVID-19 on the availability of staff responsible for AMR activities Ability to work with existing AMR partnerships, Availability of public health staff to respond to e.g. international, regional laboratory or 67% routine AMR activities, e.g. reporting, outbreak response, including healthcare associated infections, foodborne/enteric infections, facility networks (n=72) sexually transmitted diseases (n=67) Ability to create new AMR partnerships, e.g. Availability of medical doctors for AMR international, regional laboratory or facility 60% 25% activities, e.g. stewardship, infection 69% 21% networks (n=73) prevention and control (n=68) **Economic** Oversight and accountability by national AMR Availability of nursing staff for AMR activities, coordinating body of ongoing AMR activities 63% 30% e.g. stewardship, infection prevention and 71% control (n=68) Availability of infection control focal persons for AMR activities (n=68) 29% c. Reported impact of COVID-19 on diagnostics and laboratory testing for AMR Availability of environmental/cleaning service workers (n=67) Number of clinical cultures, i.e. workload of routine microbiology (culture, susceptibility 67% 26% **Feasible** testing) (n=69) Availability of laboratory staff for AMR 50% Number of screening cultures to detect multidrug 65% diagnostics and testing (n=68) resistant organisms (n=68) f. Reported impact of COVID-19 on AMR data information systems + Turn-around time of antimicrobial susceptibility results (n=70) 60% Changes to procedures and infrastructure of laboratory information systems for AMR reporting 17% Realistic Ability to carry out routine laboratory quality management activities (n=68) 53% Ability to carry out molecular testing, including Changes to procedures and infrastructure of Whole Genome Sequencing, for multidrug resistant 46% hospital clinical information systems for AMR 15% Ability to provide training for laboratory personnel (n=69) Proportion of countries reporting: Moderate change No change Large change **Ethical** φ 500 100

7-Jan-20 7-Feb-20 7-Mar-20 7-Apr-20 7-May-20 7-Jun-20 7-Jul-20

Resilient

7-Jan-21

7-Feb-21

7-Aug-20 7-Sep-20 7-Oct-20 7-Nov-20 7-Dec-20

What

Individual

Ecological

PID: 0000001

Date: 05-06-2024

Organism: 80166006 Streptococcus pyogenes (organism)

PID: 0000001

Date: 06-06-2024

Observation: 267102003 Sore throat (finding)

Post code: W12 xx

Postcode: W12 0NN

LOSA (lower layer super output area): E01001876

Multiple index of deprivation: 2

Deprivation rank: 5,098 (out of 32,844)

What

Individual

Deterministic:

 Exact one-to-one character matching using unique identifier universal across datasets

Probabilistic:

- Matching cases / data entries based on multiple non-unique variables.
- A hierarchy of variables can be defined based on data quality of each variable (e.g., admission date > DOB > gender > postcode).
- A distance between two datasets can be calculated between complete agreement and complete disagreement.

What

De-duplication: define rule to de-duplicate when the data rows are not identical (*e.g.*, the worst-case scenario is taken when there are multiple susceptibility test results for one isolate).

AST test level: R + I + I = R, I + S = I, I + undetermined = I

Isolate level: to determine whether it is a contaminant (blood commensals)

Patient level: (common practice) multiple blood specimens were taken on the same date these were combined and account for 1 test only, subsequent patient specimens were excluded

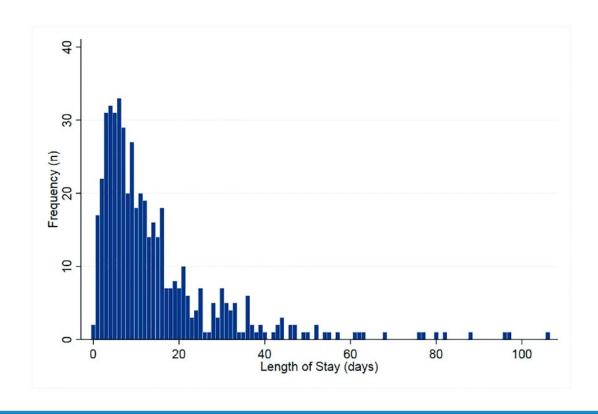
Convert to infection episode (case): based on pre-defined time window, determined by incubation period

Common bacteria: 14 days (static / rolling)

SARS-CoV-2: 60 days

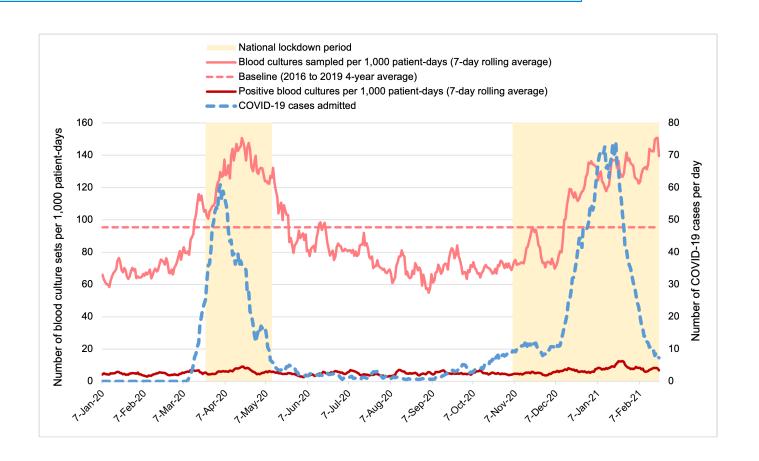
What

Imputation: replace missing values with substituted values (based on distribution). Whether there is a pattern of missing data might indicate surveillance gaps.



What

Denominator: essential when estimating prevalence and incidence, also indicates capacity / interruption of service





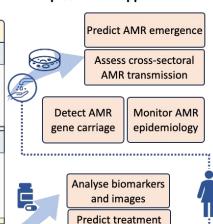


A. Different types of data available from healthcare and non-healthcare sources

Non-healthcare Healthcare Laboratory & diagnostic data Genomic data Community- Phenotypic data based Transcriptomics & metabolomics screening Diagnostic & screening Administrative & clinical data of individuals & populations Administrative Antimicrobial usage records Demographic & co-morbidities Behavioural · Clinical features & outcomes Patient & population movement factors Contextual determinants of organisations & systems Social, cultural, Policies & guidance ecological · Staffing & resources determinants Medical waste management · Health economics & financing **Environmental &** · Governance & regulation Infrastructure & technology animal sector B. Advances in data technologies ML-based clinical decision-support systems Movement & Mathematical & Automated transmission epidemiological imagine analysis simulation modelling Cloud-based computing and Trusted research data sharing environments

Data linkage and integration

C. Examples of data application



outcomes



Support therapeutic decisions

Guide IPC

Build networks of patient movements

Detect outbreaks



Design and evaluate public health policies

Evaluate health and economic impact

Support health systems strengthening

D. Data-driven outputs

Antimicrobial optimisation

- Clinical decision support systems
- Optimised therapeutics and precision medicine
- Enhanced antimicrobial stewardship

Infection prevention & control

- Prediction and prevention HCAIs
- Targeted prophylaxis
- · Patient cohorting and IPC intervention
- Enhanced environmental cleaning practices

Surveillance and tracking

- · Genomic and pathogen integrated surveillance
- Outbreak forecasting and detection tool
- Patient movement and contact tracing system

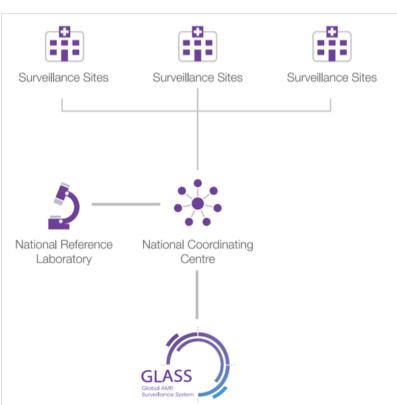
Health systems strengthening

- · Reduced health and economic burden of AMR
- Interventions for low-resourced settings to minimise inequalities
- Protected healthcare workforce



Protect individuals, populations and health systems

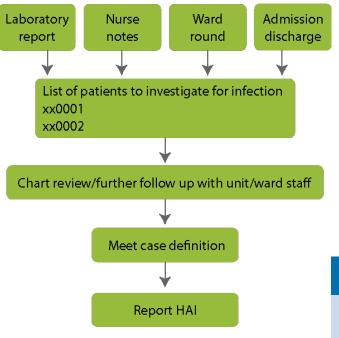
Specimen-based AMR surveillance



Key issues:

- Metadata: minimal requirement (species, AST, type of specimen); additional: demographic, epidemiological variables
- Microbiological investigation vs routine surveillance
- Isolate-based (positives only) vs specimen-based (with denominators)

Case-based HCAI surveillance



Key issues:

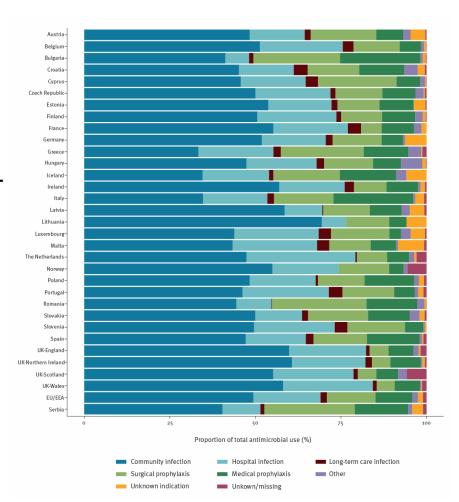
- Metadata: patient demographic, device / procedures
- Denominators: number of patients, patient-days, device-days, and procedures
- Reporting: incidence vs prevalence
- Risk stratification
- Case identified with or without individual review

Review	No review
IPC workload – disruption Delay / missed in reporting	Can be retrospectively done with routine data but with limited validity
Risk factors / risk stratifications Treatment and outcomes	

AMU surveillance

Key issues:

- Prescribing indicators relying on pointprevalence surveys
- Prescribing vs dispensing
- Over-the-counter / online purchase
- Data from health and/or non-health sectors (e.g., importing/exporting, retailers, pharmaceutical firms)



Genomic surveillance

Dr Patrick Musicha (Malawi)





Why within CAMO-Net

Mutual goal

What is available

Gaps, limitations, and assumptions

Approaches that tolerate the imperfection and get the job done

Feasible for all countries / settings without asking for more

Mutual challenges

Overview

- Welcome
- Data linkage: why, how, what to achieve within CAMO-Net
- Ongoing work in CAMO-Net UK
- Activity



NHS Discover-NOW



Imperial College Healthcare

NHS Trust

CARE



AMR & pathology

Routine EHR (secondary care)

Routine HER (primary care)

Administrative records





North West **London Pathology**

PK / TDM

- Continued and enhanced monitoring of HCAI using diagnosis-based routine EHR
- Estimating prescribing and other economic cost associated with HCAI
- Improving case identification and outcome assessment in primary care
- Addressing social determinants

HCAI: diagnosis-based EHR

Parents

- Complication of urinary catheter (disorder)
- Infection associated with catheter (disorder)
- Infection associated with genitourinary device (disorder)
- Infection following procedure (disorder)
- Urinary tract infectious disease (disorder)

Urinary tract infection associated with catheter (disorder)

SCTID: 700372006

700372006 | Urinary tract infection associated with catheter (disorder) |

Catheter-associated urinary tract infection

Urinary tract infection associated with catheter

Urinary tract infection associated with catheter (disorder)

Pathological process → Infectious process

Finding site → Urinary system

After → Catheterization of urinary

Associated with \rightarrow Urinary catheter, device

Children (1)

Infection of bladder catheter (disorder)

Parents

- Infection associated with catheter (disorder)
- Infection associated with vascular device (disorder)
- Infection following procedure (disorder)
- Infection of bloodstream (disorder)
- > Infectious disease of cardiovascular system (disorder)

■ Infection of bloodstream cooccurrent and due to central venous catheter in situ (disorder)

SCTID: 736152001

736152001 | Infection of bloodstream co-occurrent and due to central venous catheter in situ (disorder) |

Infection of bloodstream cooccurrent and due to central venous catheter in situ (disorder)

Infection of bloodstream cooccurrent and due to central venous catheter in situ

Central venous catheter associated bloodstream infection

CLABSI - central line associated bloodstream infection

After → Central venous cannula insertion

Causative agent \rightarrow Central venous catheter, device

Pathological process → Infectious process

Finding site → Structure of cardiovascular system

HCAI: diagnosis-based EHR

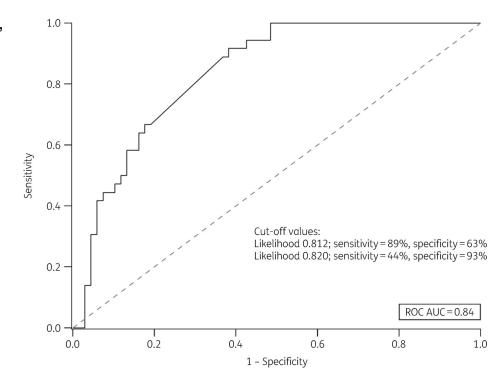
Challenges / scope of work:

- Correctly identify HCAI cases
- Use admission dates / previous admissions to determine healthcare association
- Use procedures (e.g., insertion of central venous catheter) to identify device-associated infections – removal does not get documented
- Use free-text notes (when available) to validate case identification

HCAI: diagnosis-based EHR with pathology

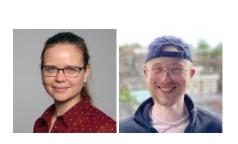
- Six blood parameters (CRP, WBC, bilirubin, creatinine, ALT and alkaline phosphatase)
 were used in the model.
- Mean likelihood estimates for those with and without infection were significantly different. The infection group had a likelihood of 0.80 and the non-infection group 0.50.

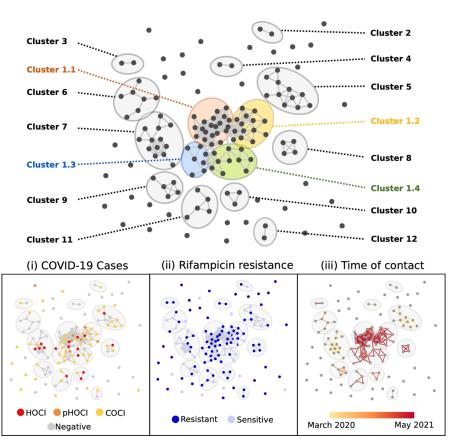




HCAI: diagnosis-based EHR with microbiology

- All 17 BSI cases would be considered blood contaminants if following the current case criteria.
- Reconstruction of patient contacts can be used to identify outbreaks of different pathogens.





HCAI: diagnosis-based EHR

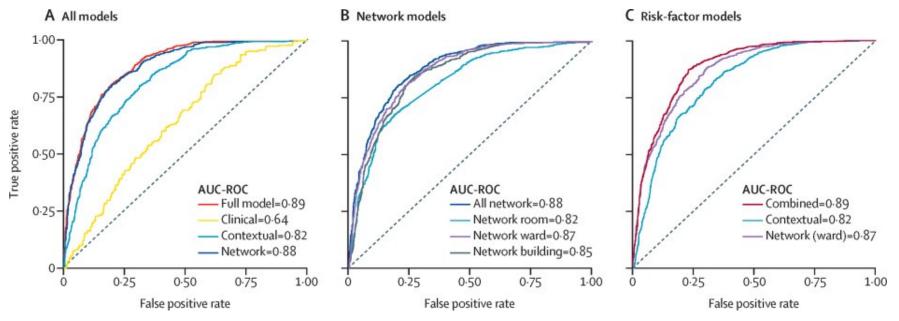
Challenges / scope of work:

- Risk prediction using pathology data: infection acquisition, treatment response
- Re-visit case criteria to determine true infections / clinically relevant infections

HCAI: risk prediction

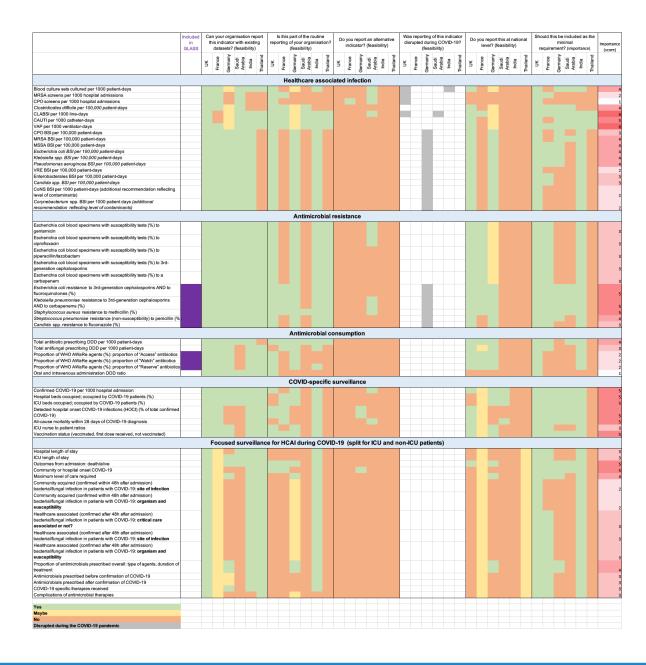


- The framework was highly predictive across test data with all variable types (AUC-ROC 0.89) and similarly predictive using only contact-network variables (0.88).
 Prediction was reduced when using only hospital contextual (0.82) or patient clinical (0.64) variables.
- A model with only 3 variables (network closeness, direct contacts with infectious patients, and hospital prevalence) achieved AUC-ROC of 0.85.

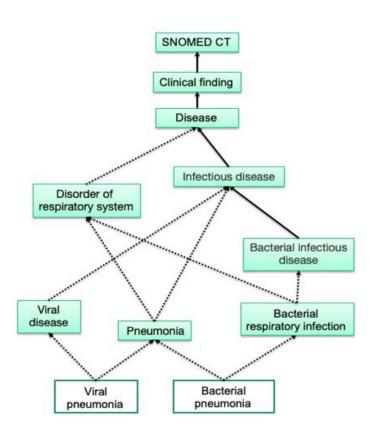


Enhance HCAI surveillance and risk prediction

- Standardisation
- Expansion



Community: case identification



Tiered approach based on what is commonly included in primary care records:

- Coded diagnosis
- Diagnosis derived from tests
- Observations / complaints
- Prescriptions

Community: case identification

Challenges / scope of work:

- Correctly identify community-acquired cases
- Look for either order of tests or test results to confirm whether diagnostics were performed
- Use antibiotic history and hospital admission to determine healthcare association

Across all sectors: address health inequalities

Challenges / scope of work:

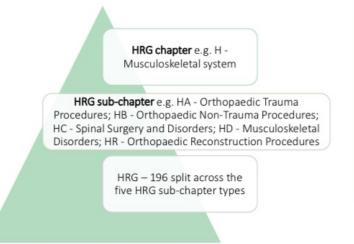
- people with multiple long-term health conditions and polypharmacy
- young carers, looked after children or care leavers
- care home residents
- residents in overcrowded households
- inclusion health groups (homeless, drug and alcohol dependent, Roma / Gypsy / traveller communities, sex workers, migrants, people in contact with justice systems and victims of modern slavery) - normally with limited data

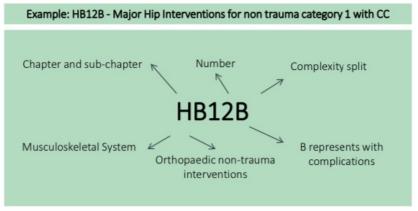
Across all sectors: the economic burden of AMR

- Item-based micro-costing
- Healthcare resource group-based micro-costing

What is Healthcare Resource Group

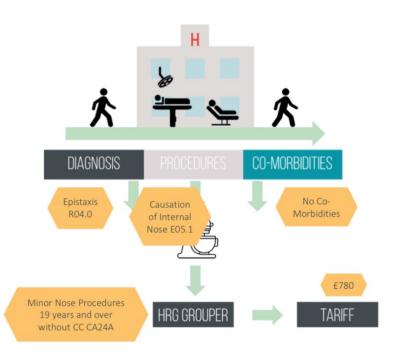
 A HRG is a set of diagnoses or procedures which are similar in terms of care delivered and resource use





In order to reflect the complexity of care delivered, HRGs capture: (i) comorbidities (ii) complications (iii) age (d) length of stay.

How a tariff is set



- Cost: national average
- Currency
- Efficiency: to adjust for the gap between commissioner funding allocation and supplier cost

Across all sectors: the economic burden of AMR

Challenges / scope of work:

- consistently defined comparators: AMR vs no AMR
- elements / items to cost
- time horizon
- in-direct medical costs
- loss of medical workforce due to transmission.

Messages

- Data linkage and integration is the way forward.
- At CAMO-Net UK, our goal is to develop feasible, contextually fit
 methods to support curation, processing, linkage, and analysis of data
 for all partner institutes and countries.
- Technological and ethical considerations? Our activity

Overview

- Welcome
- Data linkage: why, how, what to achieve within CAMO-Net
- Ongoing work in CAMO-Net UK
- Activity

Activity

- What data do you have in primary and secondary care
- Level of linkage: within one hospital, multiple hospitals, hospitals and primary care)
- Coverage: regional / sub-national, national, international
- Types of data: diagnoses, laboratory, prescribing
- Data pre-processing: clinical vocabulary / codes vs free-text
- Who are the data controllers
- Governance mechanism