# **Queen’s Doctoral Training Programme - Multi-dimensional approaches to understanding microbe/host interactions in the context of disease, therapeutics and community resilience**

|  |  |
| --- | --- |
| **Title of studentship** | Determination of appropriate antibiotic use and risk prediction of developing multidrug-resistant Gram-negative bacteria: A Machine Learning approach |
| **Project summary (max 250 words – this will be used to advertise the project if selected).** | It has been estimated that bloodstream infection (BSI) affects approximately 30 million people with 6 million deaths globally. Gram-negative bacteria *(e.g. Klebsiella pneumoniae, E. coli)* BSIare associated with more than 30% of hospital-acquired infections. Despite multi-resistant Gram-negative bacteria being one of the greatest threats to global public health, there have been no recent developments in new antibiotics for Gram-negative bacteria. The 2018 World Health Organisation Global Antimicrobial Resistance and Use Surveillance System (GLASS) report showed the global antimicrobial resistance (AMR) varies significantly within and between country. The GLASS report also highlighted many antibiotics are now less effective for treatment. Prescribing appropriate antibiotics plays a key role in tackling emerging AMR. The threat of accelerating AMR has been a major concern in the post-COVID era. The application of machine learning (ML) techniques has been successfully applied to predict diagnosis, outcomes, and disease progression. By applying ML techniques, some clinical factors have been identified to be associated with the risk of developing resistance to antibiotics. However, most of these associations have been identified from small patient cohorts and/or single hospital. The overall aim of this project is to apply ML algorithms to determine appropriate empirical antibiotic use and predict the risk of developing resistance in patients with Gram-negative BSI using two national hospital-based databases in Northern Ireland and Hong Kong. Utilising two large national datasets will provide an international comparison and enable the development of accurate ML prediction models on empirical antibiotic selection for Gram-negative BSI treatment. |
| **Supervisor(s)** | Dr Yingfen Hsia (School of Pharmacy), Professor Carmel Hughes (School of Pharmacy), Dr Deepak Padmanabhan (School of Electronics, Electrical Engineering and Computer Science) |
| **What types of new collaborative relationships would this studentship support (e.g. development of national and international collaborations or industrial involvement/financial support) (100 words max)** | This is a collaborative project between QUB Schools of Pharmacy and Electronics, Electrical Engineering and Computer Science, the Northern Ireland Public Health Agency (PHA) and the University of Hong Kong Department of Pharmacology and Pharmacy. Through this new and unique collaboration, we will be able to access national data on antibiotic use and antibiogram data for patients with Gram-negative bloodstream infections during their hospital stay in Northern Ireland and Hong Kong.  We envisage this new collaboration will support development of new projects in infectious disease research. |
| **Research centre / School** | Pharmacy/ Electronics, Electrical Engineering and Computer Science |
| **Subject area** | Pharmaco-epidemiology, antimicrobial resistance, machine learning |
| **Candidate requirements / Key skills required for the post.** Please note for the QUB-DTP awards applicants must have a 1st or 2.1 Honours degree (or equivalent) in a relevant subject. | Applicants should have a 1st or 2.1 honours degree (or equivalent) in a relevant subject including Pharmacy, Pharmaceutical Sciences, Statistics, Mathematics, Computer Science or a closely related discipline. Students who have a 2.2 honours degree and a Master’s degree may also be considered, but the School reserves the right to shortlist for interview only those applicants who have demonstrated high academic attainment to date. |
| **Relevant links for project advertisement/ more information** | https://www.qub.ac.uk/schools/SchoolofPharmacy/Research/ResearchThemes/InfectionandAntimicrobialResistance/ |
| **Keywords for search filters** | Antimicrobial resistance, antibiotic use, machine learning |
| **Training provided through the research project** | The student will receive training on quantitative methods, systematic reviewing and meta-analysis, pharmaco-epidemiological methods, generic skills in writing and presentations skills, critical thinking, and project management. Training will also be provided on computer languages (i.e. Python) and statistical software (i.e. STATA, R, Rstudio). The student will work with a large interdisciplinary team which is highly experienced in such studies. |
| **Expected impact activities** | The findings from this project can help us to implement ML algorithms in hospital settings to prescribe appropriate antibiotics and predict antimicrobial patterns. The training provided to the student will also provide an excellent grounding for a career in medical research or healthcare information technology. |

--------------------------------------------------------------------------------------------------------------------------------------

**Project details including expertise in area and workplan (max 1200 words)**

**Supervisory team and Collaborators**

This supervisory team and collaborators have been assembled to ensure the PhD student will be supported during the project by national and international experts to provide guidance at all stages.

|  |  |
| --- | --- |
| **Academic investigator** | **Expertise** |
| Dr Yingfen Hsia  (main supervisor) | Dr Hsia is an academic pharmacist and specialises in pharmacoepidemiology. She has significant experience in antimicrobial resistance research using national and global ‘big’ databases. |
| Prof Carmel Hughes  (co-supervisor) | Prof Hughes is an academic pharmacist. She has expertise in prescribing research and has undertaken studies in antibiotic use in care homes for older people. She has developed and implemented interventions to change antibiotic prescribing in primary and long-term care, and most recently, developed a Core Outcome Set (COS) for implementation in antimicrobial stewardship studies. |
| Dr Deepak Padmanabhan  (co-supervisor) | Dr Padmanabhan is a machine learning researcher and academic. He has extensive experience in machine learning and has authored 90+ papers, 12 patents and 3 books. He is currently supervising, along with Dr Hsia, a PhD student on applying machine learning for anti-microbial resistance. |
| Prof Ian Wong  (collaborator) | Prof Wong is an academic pharmacist with expertise in pharmaco-epidemiology and ‘big data’ research. He is the head of Department of Pharmacology & Pharmacy in the University of Hong Kong and Co-Director of the Centre for Medication Optimisation Research and Education at UCL Hospital and UCL School of Pharmacy. |
| Dr Lynsey Patterson  (collaborator) | Dr Patterson is the Head of Health Protection Surveillance in PHA. She is an epidemiological scientist with responsibility for management of the surveillance of healthcare-associated infections and antimicrobial resistance (AMR). |
| Dr Muhammad Sartaj  (collaborator) | Dr Sartaj is a PHA consultant in public health and health protection. |

**Project Plan**

**Background**

It has been estimated that bloodstream infection (BSI) affects approximately 30 million people with 6 million death globally (1). Gram-negative bacteria *BSI* are associated with more than 30% of hospital-acquired infections. Studies have demonstrated that appropriate empirical antibiotic for BSI can reduce morbidity and mortality (2-4). However, there is a major concern that clinicians will empirically prescribe combinations of broad-spectrum antibiotics to reduce poor outcomes. However, excessive and inappropriate empirical antibiotic prescribing will inevitably contribute to antimicrobial resistance (AMR). Monitoring appropriate antibiotic use is the core principle of antibiotic stewardship activity to tackle AMR. The current COVID10 pandemic poses a threat that could affect antibiotic stewardship activities and accelerate antimicrobial resistance. There is an urgent need to prescribe empirical antibiotic appropriately to reduce the risks of developing AMR during the post-COVID era (5).

More recently, the increasing availability of large medical databases provides an opportunity to enable better-informed clinical diagnosis and appropriate treatment. The application of machine learning (ML) techniques to these medical databases has gained popularity in AMR research. In addition, there is current interest in implementing ML models to enhance antimicrobial stewardship in hospital electronic systems. Several studies have demonstrated that ML algorithms can be used to identify risk factors associated with the development of infectious diseases, predict risk factors, empirical antibiotics, and resistance patterns. However, these studies were limited to small patient cohorts and/or a single hospital. To apply ML techniques using large national databases will ensure sufficient data to develop robust ML algorithms. The accurate ML algorithms can potentially lead to evidence-based antibiotic prescribing policies to reduce inappropriate antibiotic use in clinical practice.

**Aims and objectives**

The overall aim of this project is to apply ML algorithms using national hospital databases to determine appropriate empirical antibiotic use and predict the risk of developing resistance in patients with Gram-negative bloodstream infections, including *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *E. coli* in Northern Ireland and Hong Kong. This will be achieved through the following objectives:

1. To conduct a systematic review to identify studies which utilised ML to identify factors associated with the risk of development of resistance to antibiotics in patients with Gram-negative bloodstream infections. The risk factors identified in systematic review study will enable us to develop ML models for Object 3 and 4.
2. To characterise clinical demographics, antibiotic prescribing patterns, and antibiotic resistance patterns in patients with Gram-negative infections. This analysis will be carried out separately for Northern Ireland and Hong Kong datasets.
3. Northern Ireland hospital data: to develop ML algorithms to create a prediction model to identify appropriate empirical antibiotics for patients with Gram-negative bloodstream infections.
4. Hong Kong hospital data: to develop ML algorithms to create a prediction model to identify risk factors associated with the development of resistance in patients with Gram-negative bloodstream infections.
5. Comparative analysis: it is anticipated the results from Objectives 3 and 4 will be different from two national databases. A comparative analysis will be carried out to evaluate the similarity and differences in empirical antibiotic selection for *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *E. coli* intwo study populations.

**Data source**

**Database in Northern Ireland Public Health Agency**

The Public Health Agency (PHA) has statutory responsibility for the monitoring of infectious disease in Northern Ireland. In this function, the PHA maintains a regional laboratory information system (The Northern Ireland Laboratory Information System: NILIS) which contains data on culture-positive results for bacterial, fungal and viral samples from the five regional laboratories in NI, as well as antibiotic information. This includes bloodstream infections which have been used for surveillance of AMR since 2017. There is also a number of enhanced surveillance programmes including one for Gram-negative bacteraemia which was introduced in 2018. In addition, PHA also maintains a regional dataset for antibiotic consumption in secondary care. This information has been available since 2013 (6-7).

**Database in Hong Kong**

The Clinical Data Analysis and Reporting System (CDARS) is a database maintained by the Hong Kong Hospital Authority. Health services provided by the Hospital Authority are available to all residents (approximately >7 million population). The electronic records contain patients’ demographic information, diagnoses, prescription information, laboratory tests and admission and discharge information. The CDARS has captured data since 1993 from all public hospitals, institutions and outpatient clinics under the Hospital Authority and has been used to conduct several epidemiological studies (8-9).

**Proposed ML models**

Model development

As AMR patterns vary in different age groups, the database will first be analysed to include Gram-negative pathogens and then divided into two datasets, a dataset containing data for children (aged ≤18 years) and a dataset containing data for adults >18 years. Both datasets will be further divided to create a training dataset and a validation dataset. The training dataset for both populations will be used to create ML algorithms which will then enable us to develop multiple training models. The sensitivity and specificity of each model will be evaluated and the most successful model, or the model which predicts the antimicrobial resistance of an infection and therefore suggests the most suited antibiotic per patient will be chosen for each study population. These models will be continuously updated until they best fit the data on hand. The model development will be carried out separately for Northern Ireland and Hong Kong datasets.

Model validation

The chosen models for the paediatric dataset and the adult dataset will be validated using the validation datasets which created during model development. The datasets will contain patient demographics and antibiogram profiles for this validation process. The models will predict the resistance profile for each individual patient and the results will be compared to the measured antibiogram profiles to fairly evaluate the performance of the model and the accuracy of the predictions. The model validation will be carried out separately for Northern Ireland and Hong Kong datasets.

We anticipate that data cleaning, data manipulation, statistical analysis, and ML development in this project will be challenging. Data cleaning is an essential process of achieving good quality data to develop ML algorithm. It will take several months for the student to understand data structure and data capture in different systems, in addition to learning the programming skills required for ML. Thus, we have carefully planned this project to ensure that the student can successfully complete the required work within the 3-year studentship period.

**Impact of project**

Antibiotic stewardship is the core activity to tackle AMR. The predictive ML algorithms developed in this project can be potentially implemented as part of antibiotic stewardship and/or change antibiotic prescribing policies in hospital setting. As the AMR patterns vary between country, results from this project will also enable us to have a comparative analysis on ML algorithms in two different populations.

**Research experience and Method Training**

This project will provide opportunities for students to learn epidemiological methods, improve analytical skills, statistical modelling, and machine learning modelling. The supervisory team is multidisciplinary, with both clinical and technical expertise in pharmacoepidemiology, antibiotic stewardship, antimicrobial resistance and machine learning modelling. In addition, student will have the opportunity to work closely with the NI PHA.

**Dissemination**

It is anticipated that several papers will be published in peer-reviewed medical journals, including a systematic review, the ML model to guide antibiotic use, and risk factor prediction of developing AMR. We will also submit abstracts to international conferences for presentation.

**References**

1. Goto M, Al-Hasan MN. Overall burden of bloodstream infection and nosocomial bloodstream infection in North America and Europe. Clin Microbiol Infect. 2013 Jun;19(6):501-9.
2. Mehrad B, Clark NM, Zhanel GG, Lynch JP 3rd. Antimicrobial resistance in hospital-acquired gram-negative bacterial infections. Chest. 2015 May;147(5):1413-1421.
3. Lodise TP Jr, Patel N, Kwa A, Graves J, Furuno JP, Graffunder E, Lomaestro B, McGregor JC. Predictors of 30-day mortality among patients with Pseudomonas aeruginosa bloodstream infections: impact of delayed appropriate antibiotic selection. Antimicrob Agents Chemother. 2007 Oct; 51(10):3510-5.
4. Paul M, Shani V, Muchtar E, Kariv G, Robenshtok E, Leibovici L. Systematic review and meta-analysis of the efficacy of appropriate empiric antibiotic therapy for sepsis. Antimicrob Agents Chemother. 2010 Nov;54(11):4851-63.
5. Getahun H, Smith I, Trivedi K, Paulin S, Balkhy HH. Tackling antimicrobial resistance in the COVID-19 pandemic. Bull World Health Organ. 2020 Jul 1;98(7):442-442A.
6. Patterson L and Bradley D. HSC Public Health Agency: surveillance of antimicrobial use and resistance in Northern Ireland, Annual report 2017. Available: <https://www.publichealth.hscni.net/sites/default/files/2019-02/AMR%20annual%20report%202017.pdf> (access on Nov 2020).
7. Patterson L, Nugent C, Sartaj M. HSC Public Health Agency: surveillance of antimicrobial use and resistance in Northern Ireland, Annual report 2018. Available: <https://www.publichealth.hscni.net/sites/default/files/2019-02/AMR%20annual%20report%20final%202018.pdf> (access on Nov 2020)
8. Wong AY, Wong IC, Chui CS, Lee EH, Chang WC, Chen EY, Leung WK, Chan EW. Association Between Acute Neuropsychiatric Events and Helicobacter pylori Therapy Containing Clarithromycin. JAMA Intern Med. 2016 Jun 1;176(6):828-34
9. Lao KSJ, Wong AYS, Wong ICK, Besag FMC, Chang WC, Lee EHM, Chen EYH, Blais JE, Chan EW. Mortality Risk Associated with Haloperidol Use Compared with Other Antipsychotics: An 11-Year Population-Based Propensity-Score-Matched Cohort Study. CNS Drugs. 2020 Feb;34(2):197-206.