

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

*Data-driven web-based intelligent decision support
system for infection management at point of care*

Author: Bernard Hernandez Perez

Supervisor: Dr. Pantelis Georgiou

Co-Supervisor: Dr. Pau Herrero

Internal Examiner: Dr. Alex Bottle

External Examiner: Prof. Marc Mendelson

I hereby declare that this dissertation and the work described in it are my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Copyright © 2018 by Bernard Hernandez Perez: *Data-driven web-based intelligent decision support system for infection management at point of care*

To my father
– this thesis is as much yours as it is mine

"A good decision is based on knowledge." – Plato

Socrates: What, Lysimachus, are you going to accept the opinion of the majority?

Lysimachus: Why, yes, Socrates; what else am I to do?

Socrates: And would you do so too, Melesias? If you were deliberating about the gymnastic training of your son, would you follow the advice of the majority of us, or the opinion of the one who had been trained and exercised under a skilful master?

Melesias: The latter, Socrates; as would surely be reasonable.

Socrates: His one vote would be worth more than the vote of all us four?

Melesias: Certainly.

Socrates: And for this reason, as I imagine, –because a good decision is based on knowledge and not on numbers?

Melesias: To be sure.

Socrates: Must we not then first of all ask, whether there is any one of us who has knowledge of that about which we are deliberating? If there is, let us take his advice, though he be one only, and not mind the rest; if there is not, let us seek further counsel.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to Dr Pantelis Georgiou for giving me the opportunity, support and freedom to pursue this research. The confidence you put in me was bracing.

I would also like to show my appreciation to Dr Pau Herrero and Dr Timothy Rawson for their knowledge and guidance through this journey. Also, many thanks to Prof Alison Holmes and Prof Marc Mendelson for facilitating me the opportunity to broaden my research experience in one of the most beautiful places in the world.

To the people in B422, I blame all of you. Writing this thesis has been an exercise in sustained suffering. For those of you who have played the larger role in prolonging my agonies with your encouragement, support and friendship well... you know who you are and I am endlessly grateful. In particular, to Amparo Güemes who not only carefully reviewed this thesis but also pushed me through with all her love and support. I am very much looking forward to our next adventure!

Last, my deep and sincere gratitude to my family. I am forever indebted to you for giving me the opportunities and experiences that have made me who I am. You selflessly encouraged me to explore new directions in life and seek my own destiny. This journey would not have been possible if not for you, and I dedicate this milestone to you.

ABSTRACT

Infectious diseases are caused by the invasion of pathogenic microorganisms such as bacteria, viruses or fungi and are one of the leading causes of mortality worldwide. In the last years, there has been a significant increase in the ability of these microorganisms to resist antimicrobials which were previously effective. This phenomenon, denoted as antimicrobial resistance (AMR), has become a noticeable obstacle to treat infections in health care with misuse and overuse of antimicrobials as one of the leading drivers.

This thesis presents a novel clinical decision support system for infection management to provide personalized, accurate and effective diagnostics at point of care. The proposed system, which has been denoted as EPiC IMPOC, incorporates two main decision support engines: case-based reasoning to facilitate vital sign collection, patient monitoring and further inspection of past similar cases and probabilistic inference to provide stepwise guidance within the infection management pathway followed by clinicians. In addition, a number of local AMR statistics are automatically computed from susceptibility test data to promote education and awareness among clinicians. This decision support system has been implemented as a web-based platform that can be accessed at the point of care from computers or hand-held devices.

The design and implementation of the system has been performed incrementally. As such, the server has been divided into discrete and reusable modules. Firstly, the AMR statistics have been computed and compared with the existing literature to better describe the scope of the problem. After this, the case-based reasoning methodology was included to inform physicians of previous past cases to assist in antimicrobial therapy selection. In order to evaluate the validity and usability of the system, a pilot study was conducted in the ICU. The system advocated the same results as those suggested by infection specialists in 84% of the cases. Furthermore, participants highlighted the utility of the case-based reasoning engine to promote knowledge transfer among health care professionals and the benefits of having access to real-time patient data at the point of care. The system usability score was 68.5 which is above average. On the other hand, participants suggested that the provision of more specific support would be beneficial. For this reason, the probabilistic inference module was included into the system to provide the likelihood of positive culture (AUCROC over 0.90) and the most plausible sites of infection (AUCROC within the range 0.82–0.96). The translational utility of the generated predictive models was assessed retrospectively considering also scenarios with missing data and imbalanced classes.

All these elements combined result in a state-of-the-art clinical decision support system which assists physicians on multiple areas within infection management to facilitate the provision of evidence-based and personalized medicine.

CONTENTS

Acknowledgements	vii
1 Introduction	1
1.1 Motivation	1
1.2 Research objectives	2
1.3 Thesis organisation	3
1.3.1 Chapter 2: Background	3
1.3.2 Chapter 3: CDSS design and implementation	3
1.3.3 Chapter 4: Antimicrobial resistance surveillance	3
1.3.4 Chapter 5: Case-based reasoning	3
1.3.5 Chapter 6: Probabilistic inference	4
1.3.6 Chapter 7: Conclusions	4
2 Background	9
2.1 Infectious diseases	9
2.1.1 Infectious microorganisms	9
2.1.2 Site of infection	10
2.1.3 Modes of transmission of infectious diseases	11
2.2 The burden of antimicrobial resistance	11
2.2.1 What are the main drivers of AMR?	11
2.2.2 What are the consequences of AMR?	13
2.2.3 How can we mitigate AMR?	13
2.3 Clinical decision support systems	14
2.3.1 Overview of decision support engines	14
2.3.2 CDSSs for infection management	15
2.3.3 Methods for infectious disease diagnosis	17
2.4 Towards comprehensive decision support	19
2.4.1 The five ‘rights’ of a CDSS	19
2.4.2 The desired characteristics of a CDSS	19
2.4.3 The proposed solution	21
3 CDSS design and implementation	31
3.1 Decision support system overview	31
3.2 Part I: Outline of decision support modules	33
3.3 Part II: The graphical user interface	34
3.3.1 User interface design strategies	34
3.3.2 Overview of the design process	35

3.3.3	Illustrating the medical record of a patient	37
3.3.4	Understanding the main control panel	38
3.3.5	Design techniques and usability principles within EPiC IMPOC	40
3.3.6	The patient engagement module	42
3.3.7	The AMR surveillance module	44
3.4	Discussion	46
3.4.1	Providing clinical training to facilitate adoption	47
3.4.2	Limitations	47
3.5	Conclusion	48
4	Antimicrobial resistance surveillance	53
4.1	How to measure AMR?	53
4.2	Materials and methods	55
4.2.1	Susceptibility test data	55
4.2.2	Generation of resistance time series signals	55
4.2.3	Trend and stationarity in time series	56
4.2.4	Pearson correlation coefficient	57
4.2.5	Sigmoid function	57
4.2.6	Statistical significance among regression methods	57
4.2.7	Outline of the experiments conducted	58
4.3	Experiment I: Overview of susceptibility test data	59
4.4	Experiment II: Antimicrobial resistance rates	60
4.5	Experiment III: Antimicrobial spectrum of activity	62
4.5.1	The antimicrobial spectrum of activity index	62
4.5.2	Spectrum of activity summary for urine culture	63
4.5.3	Enhancing traditional antimicrobial spectrum categorization	64
4.6	Experiment IV: Antimicrobial resistance trends	66
4.6.1	Regression analysis for trend estimation	66
4.6.2	Analysis of the robustness of the methods	67
4.6.3	<i>Escherichia coli</i> : AMR in urine cultures	69
4.6.4	<i>Escherichia coli</i> : AMR in blood cultures	70
4.6.5	<i>Staphylococcus aureus</i> : AMR in wound cultures	71
4.7	Discussion	72
4.7.1	Susceptibility testing: behaviour and guidelines	72
4.7.2	Advantages of overlapping time intervals in surveillance	72
4.7.3	Regression analysis: benefits and drawbacks	72
4.7.4	The importance of surveillance data	73
4.7.5	Limitations	73
4.8	Conclusions	74
5	Case-based reasoning	79
5.1	The case-based reasoning methodology	79
5.1.1	The use of prototypes	80
5.2	Materials and methods	81
5.2.1	Definition of generic distance metrics	81
5.2.2	Definition of case features: antimicrobial therapy selection in the ICU	82

5.2.3	Definition of similarity metrics: antimicrobial therapy selection in the ICU	83
5.2.4	Assigning feature importance	85
5.2.5	Performance metrics for information retrieval problems	86
5.2.6	Performance metrics for classification problems	87
5.2.7	Overall performance of the CBR model	88
5.2.8	Statistical analysis	88
5.2.9	Outline of the experiments conducted	88
5.3	Experiment I: A simple scenario (iris flower dataset)	89
5.3.1	Distributions of the iris flower dataset features	89
5.3.2	Analysis of the CBR retrieval performance	89
5.4	Experiment II: Imbalanced categories (ecoli protein location dataset)	92
5.4.1	Distributions of the ecoli protein location dataset features	92
5.4.2	Analysis of the CBR retrieval performance	93
5.5	Experiment III: Intensive Care Unit	96
5.5.1	Pilot study by infection specialists	96
5.5.2	Usability study of the CDSS	97
5.6	Discussion	99
5.6.1	Benefits and limitations of CBR	99
5.6.2	Feedback provided from users	100
5.6.3	The infection management pathway	101
5.6.4	Limitations	101
5.7	Conclusions	102
6	Probabilistic inference	107
6.1	Brief introduction to supervised machine learning	107
6.2	Materials and methods	109
6.2.1	Selected pathology biochemical markers	110
6.2.2	Assembling data for inference	110
6.2.3	Challenges in clinical data: preprocessing	111
6.2.4	Model evaluation	112
6.2.5	Model calibration	113
6.2.6	Statistical analysis	113
6.2.7	Outline of the experiments conducted	113
6.3	Data Analysis	114
6.3.1	Laboratory tests frequency	114
6.3.2	Profile completeness	115
6.4	Experiment I: Prediction of positive culture	116
6.4.1	Distributions of the selected biochemical markers	116
6.4.2	Infection risk inference on complete profiles	116
6.5	Experiment II: Prediction of site of culture	119
6.5.1	Distribution of the selected biochemical markers	119
6.5.2	Site of infection risk inference on complete profiles	120
6.6	Experiment III: Voting prediction of positive culture	122
6.7	Experiment IV: Prediction of the gram status of bacteria	123
6.8	Experiment V: Missing data and imbalance classes	123
6.8.1	Inference on scenarios with missing data	124

6.8.2	Inference on scenarios with class imbalance	124
6.8.3	Probability calibration and thresholds	124
6.9	Discussion	126
6.9.1	Selecting suitable biochemical markers	126
6.9.2	Addressing class imbalance	126
6.9.3	Effect of missing inputs in prediction	126
6.9.4	Selecting a suitable algorithm	127
6.9.5	Translational utility	128
6.9.6	Limitations	128
6.10	Conclusions	128
7	Conclusions and future perspective	135
7.1	Satisfying the needs in infection management	135
7.1.1	Need of integration with electronic health records	135
7.1.2	Need of a point-of-care decision support system	136
7.1.3	Need of understanding local AMR patterns	136
7.1.4	Need to align with the infection management pathway	136
7.2	Overall discussion and contributions	138
7.3	Future research directions	143
7.3.1	Effective communication of local AMR patterns	143
7.3.2	Improving history review and case retrieval	143
7.3.3	Enhancing stepwise decision support in infection management	144
7.3.4	Integration of further technology	145
7.3.5	Implementation in low- and middle-income countries	145
7.3.6	Summary of potential research directions	146
7.4	Final conclusion	146
A	List of publications	i
A.1	Journal articles	i
A.2	Conference: abstracts and posters	ii
A.3	Conference: oral presentations	ii
B	Literature Review	iii
B.1	Clinical decision support systems for infection management	iii
B.2	Diagnosis of infectious diseases using machine learning	v
C	Documentation	xv
D	List of acronyms: organisms and antimicrobials	xvii
E	Definition of the attributes	xxv

LIST OF FIGURES

3.1	EPiC IMPOC: high-level system architecture diagram	32
3.2	EPiC IMPOC main panel: basic interaction diagram	38
4.1	Susceptibility test record attributes	55
4.2	Sigmoid function examples	57
4.3	Graphical description of the experiments	58
4.4	Proportion of sample types (cultures)	59
4.5	Experiment I: antimicrobial resistance rates in urine cultures (clustermmap) . .	61
4.6	Experiment II: antimicrobial spectrum of activity: graphical description . . .	63
4.7	Experiment III: high-level methodology diagram for antimicrobial resistance trend estimation	66
4.8	Experiment III: distribution of paired SART distances.	68
4.9	Experiment III: AMR summary for <i>Escherichia coli</i> in urine samples	69
4.10	Experiment III: AMR summary for <i>Escherichia coli</i> in blood samples	70
4.11	Experiment III: AMR summary for <i>Staphylococcus aureus</i> in wound samples .	71
5.1	The CBR cycle	80
5.2	High-level methodology diagram for model creation and evaluation.	81
5.3	Experiment I: distributions for iris categories	89
5.4	Experiment I: nDCG comparison	91
5.5	Experiment II: distributions for ecoli protein location categories	93
5.6	Experiment II: nDCG comparison	95
6.1	High-level methodology diagram for model creation and evaluation.	109
6.2	Biochemical markers: daily profile	111
6.3	Graphical description of the experiments	114
6.4	Biochemical markers: frequency and completeness	115
6.5	Biochemical markers: distributions for C- and C+	116
6.6	Positive culture inference: ROC, PR and calibration curves.	118
6.7	Biochemical markers: distributions for C- and sites	119
6.8	SVM-based estimators: ROC and PR curves (site).	121
6.9	ANN-based estimators: ROC and PR curves (site).	121
6.10	Summary of estimators: ROC, PR, calibration and threshold curves	125
7.1	Acute infection management pathway	137
7.2	EPiC IMPOC: graphical summary	142

LIST OF TABLES

1.1	Research questions	2
2.1	Leading drivers of AMR	12
2.2	Literature review: comparison of CDSSs for infection management	18
2.3	Literature review: comparison of methodologies for infection diagnosis	18
2.4	The five rights of a CDSS: brief description and examples	19
2.5	CDSS for infection management: desired attributes	20
2.6	CDSS for infection management: desired functionality	20
3.1	EPiC IMPOC: summary of the design proces	35
3.2	EPiC IMPOC: the medical record of a patient	37
3.3	EPiC IMPOC: the main control panel	39
3.4	EPiC IMPOC: design techniques and usability principles	41
3.5	EPiC IMPOC: the patient engagement report	43
3.6	EPiC IMPOC: the AMR surveillance module	45
4.1	Resistance indexes: description and formula	54
4.2	Resistance time series: description of generation strategies	56
4.3	Experiment I: top 10 pathogens in urine cultures	59
4.4	Experiment I: top 10 pathogens in wound cultures	59
4.5	Experiment I: top ten pathogens in sputum cultures	59
4.6	Experiment I: top ten pathogens in blood cultures	59
4.7	Experiment II: antimicrobial spectrum of activity summary in urine cultures	64
4.8	Experiment III: AMR summary for <i>Escherichia coli</i> in urine samples	69
4.9	Experiment III: AMR summary for <i>Escherichia coli</i> in blood samples	70
4.10	Experiment III: AMR summary for <i>Staphylococcus aureus</i> in wound samples	71
5.1	Distance metrics: description and equation	82
5.2	Description of strategies to assign feature weights	86
5.3	Evaluation metrics: description and equation	87
5.4	Experiment I: iris flower dataset description	89
5.5	Experiment I: comparison of distance metrics (I)	90
5.6	Experiment I: comparison of distance metrics (II)	90
5.7	Experiment I: comparison of weight configurations	90
5.8	Experiment II: <i>Escherichia coli</i> protein location dataset description	92
5.9	Experiment II: comparison of distance metrics (I)	93
5.10	Experiment II: comparison of distance metrics (II)	93
5.11	Experiment II: comparison of weight configurations	94

5.12	Pilot study in ICU: free comments	97
5.13	System Usability Scale: statements and contribution	98
6.1	Biochemical markers: description	110
6.2	Evaluation metrics: description and equation	113
6.3	Biochemical markers: frequency and completeness	115
6.4	Experiment I: comparison of sampling methods	117
6.5	Experiment I: comparison of sampling methods (IQRx1.5)	117
6.6	Experiment I: comparison culture-positive inference	118
6.7	Experiment I: comparison culture-positive inference (IQRx1.5)	118
6.8	Experiment II: comparison site of infection	120
6.9	Experiment II: comparison site of infection (IQRx1.5).	120
6.10	Experiment III: comparison culture-positive inference	122
6.11	Experiment III: comparison culture-positive inference (ensemble)	122
6.12	Experiment IV: ANN for Gram status prediction	123
6.13	Experiment IV: ANN for Gram status prediction (IQRx1.5)	123
6.14	Algorithms: summary of properties	127
7.1	Likelihood of positive culture: existing research	138
7.2	Likelihood of positive culture: this research	138
7.3	Likelihood site of culture: existing research	139
7.4	Likelihood site of culture: this research	139
7.5	Gram status of bacteria: existing research	139
7.6	Gram status of bacteria: this research	139
7.7	Potential scores for inclusion in EPiC IMPOC	144
7.8	Summary: future research directions	146
B.1	Literature review: description of columns (infection diagnosis)	iii
B.2	Literature review: clinical decision support systems	iv
B.3	Literature review: description of columns (infection diagnosis)	v
B.4	Literature review: diagnosis of bloodstream infection (I)	vi
B.5	Literature review: diagnosis of bloodstream infection (II)	vii
B.6	Literature review: diagnosis of urinary tract infection	vii
B.7	Literature review: diagnosis of respiratory tract infection	viii
B.8	Literature review: diagnosis of surgical site infection	ix
B.9	Literature review: diagnosis of culture positive	ix
C.1	EPiC IMPOC interface: cases	xv
C.2	EPiC IMPOC interface: doctors	xvi
C.3	EPiC IMPOC interface: patients	xvi
C.4	EPiC IMPOC interface: authentication	xvi
C.5	EPiC IMPOC interface: inference	xvi
C.6	EPiC IMPOC interface: others	xvi
D.1	Antibiotics: notation and description	xvii
D.1	Antibiotics: notation and description	xviii
D.1	Antibiotics: notation and description	xix
D.1	Antibiotics: notation and description	xx

D.1	Antibiotics: notation and description	xxi
D.2	Antimicrobials: notation and description	xxii
D.2	Antimicrobials: notation and description	xxiii
E.1	CBR: feature attributes	xxv

ACRONYMS

ADF	Augmented Dickey–Fuller
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
AMR	Antimicrobial Resistance
ANN	Artificial Neural Network
API	Application Programming Interface
ARIMA	Autoregressive Integrated Moving Average
AUC	Area Under the Curve
BIC	Bayesian Information Criterion
BIL	Bilirubin
BSAC	British Society Antimicrobial Chemotherapy
CBR	Cased-Based Reasoning
CDSS	Clinical Decision Support System
COPD	Chronic Obstructive Pulmonary Disease
CPN	Causal Probabilistic Network
CRE	Creatinine
CRP	C-Reactive Protein
CSS	Cascading Style Sheets
CVS	Cross Validation Set
DTC	Decision Tree Classifier
EHR	Electronic Health Records
ESBL	Extended-Spectrum Beta-Lactamases
GNB	Gaussian Naïve Bayes
GUI	Graphical User Interface
HAI	Hospital Acquired Infection
HCAI	Healthcare Associated Infection
HIV	Human Immunodeficiency Virus
HOS	Hold Out Set
HTML	Hypertext Markup Language
ICU	Intensive Care Unit
IQR	Inter Quartile Range
KPSS	Kwiatkowski–Phillips–Schmidt–Shin

LDAP	L ightweight D irectory A ccess P rotocol
LRTI	L ow R espiratory T ract I nfection
LUTI	L ower U rinary T ract I nfection
NHS	N ational H ealth S ervices
OLS	O rdinary L east S quares
PDA	P ersonal D igital A ssistant
PDF	P ortable D ocument F ormat
PhD	P hilosophiae D octor (doctor of philosophy)
PID	P ersonal I dentification N umber
PKPD	P harmacokinetic P harmacodynamic
POC	P oint O f C are
PR	P recision R ecall
REST	R epresentational S tate T ransfer
RFC	R andom F orest C lassifier
ROC	R eceiver O perating C haracteristic
RTI	R espiratory T ract I nfection
SIRS	S ystemic I nflammatory R esponse S yndrome
SOFA	S epsis-related O rgan F ailure A ssessment
SQL	S tructured Q uery L anguage
SMOTE	S ynthetic M inority O versampling T echnique
SUS	S ystem U sability S cale
SVM	S upport V ector M achine
UK	U nited K ingdom
URTI	U pper R espiratory T ract I nfection
USA	U nited S tates of A merica
UTI	U rinary T ract I nfection
UUTI	U pper U rinary T ract I nfection
WBC	W hite B lood C ell
WLS	W eighted L east S quares

INTRODUCTION

1.1 MOTIVATION

Infectious diseases are caused by the invasion of disease causing pathogens such as bacteria, viruses, parasite or fungi and are one of the leading causes of mortality worldwide [1]. Infections are frequently treated with antimicrobials which are drugs that kill or stop the growth of microbes. Since the creation of first antimicrobials, researchers and other specialists were concerned about their misuse and possible complications that could arise. Despite all their efforts to disseminate awareness, there has been a significant increase in the ability of microbes to resist antimicrobials which were previously effective [2, 3]. In the last years, antimicrobial resistance (AMR) has become a noticeable obstacle to treat infections in health care with estimates that it will be responsible for more than ten million deaths by 2050 [4]. A large body of research has identified the misuse and overuse of antimicrobials as one of the leading drivers of AMR [5] yet antimicrobials are among the most commonly prescribed drugs in human medicine [6]. One of the reasons prescribers use antimicrobials inappropriately is lack of diagnostic certainty [7].

The introduction of electronic medical records in health care has accelerated the implementation of decision support systems to improve patient outcomes by promoting the provision of evidence-based [8] and personalized [9] medicine. These systems integrate medical knowledge and patient data to support and assist physicians and have demonstrated to increase quality of care in medicine [10]. However, innovation adoption within infection management is still low [11]. In the last years, machine learning techniques have influenced an enormous number of research areas such as health care in order to automatically extract patterns from data [12]. For instance, a plethora of research has emerged to exploit electronic medical records in order to measure antimicrobial resistance rates [13–18], analyse resistance trends [19] or provide decision support for infection diagnosis [20, 21], therapy selection [22–25] or antimicrobial dosing [26]. However, little efforts have been made to integrate all this research into a comprehensive decision support system to enhance infection management.

Therefore, there is an opportunity to develop a data-driven decision support system that exploits such large amount of health data through the use machine learning techniques to provide personalized, accurate and effective support for infection management at the point of care.

1.2 RESEARCH OBJECTIVES

The research hypothesis of this PhD thesis is that an appropriately designed clinical decision support system for infection management could greatly benefit clinicians in many different aspects. Firstly, it would improve reliability and consistency of collecting vital signs; their visualization, interpretation and analysis at point of care. Secondly, it would promote continuity, interpersonal communication, education and knowledge transfer. Finally, providing advice through the infection management pathway followed by clinicians would promote the practice of evidence-based and personalized medicine. Therefore, the general question addressed by this research is:

Can an Intelligent Decision Support System be developed and exploited in clinical workplace to improve patient outcomes while promoting appropriate antibiotic prescription practices?

Subsequently, this thesis has been structured to provide some insight into the questions presented in Table 1.1 for each individual chapter.

Table 1.1: Research questions targeted within this thesis by chapter.

Chapter 2	Background	Chapter 3	System design	Chapter 4	AMR surveillance
What are the main drivers of AMR?		Are accessibility at point of care and integration into electronic health records essential in a CDSS?		What hospital-specific AMR statistics should be accessible to promote education and increase awareness among health care workers?	
Which of these drivers could be targeted through the implementation of a CDSS in health care settings? How?		What modules for decision support should be implemented within the CDSS to provide advice across infection management?		How to use microbiology data to measure and therefore effectively communicate:	
What are the goals of existing CDSS and which limitations are obstructing implementation and adoption?		How to design the user interface to provide decision support while allowing data collection and patient monitoring?		- antimicrobial resistance rates ?	
What are the essential properties of a CDSS to overcome such limitations and enhance infection management?				- antimicrobial resistance trends ?	
				- antimicrobial spectrum of activity ?	
Chapter 5	Case-based reasoning	Chapter 6	Probabilistic inference		
What methodology could be employed to retrieve previous similar cases to provide decision support?		What is the decision pathway followed by clinicians during infection management in secondary care?			
What vital signs, symptoms or laboratory tests are relevant for clinicians and therefore should be included in the clinical record of a patient?		By using a reduced set of biochemical markers, is it feasible to provide stepwise decision support to...			
What similarity measure should be defined to rank and retrieve previous patient cases?		- determine the risk of infection ?			
		- assess the site(s) of infection ?			
		- discern Gram+ / Gram- bacteria?			
		- identify the infectious organism ?			

To summarise, the aim of this research is to understand the needs of health care workers such as nurses, physicians and infection specialists to develop a fully functional point of care clinical decision support system to improve the management of infectious diseases in hospitals.

1.3 THESIS ORGANISATION

The presented thesis has been organised in six chapters. For the sake of clarity, a brief overview of the chapters is presented below.

1.3.1 Chapter 2: Background

This chapter provides a general overview on infectious diseases (section 2.1) and discusses the main drivers, challenges and consequences of antimicrobial resistance (section 2.2). Clinical decision support systems (CDSSs) are considered an effective approach to tackle most contributing drivers by improving prescription practices. For such reason, the chapter presents popular decision support engines (subsection 2.3.1), reviews the current state of CDSSs in infection management (subsection 2.3.2) and explores research trends for infectious disease diagnosis from patient data (subsection 2.3.3). To conclude, the chapter discusses the shortcomings of current CDSSs and how the research proposed in this thesis aims to overcome such limitations.

1.3.2 Chapter 3: CDSS design and implementation

This chapter provides an overview of the clinical decision support system implemented in this PhD. Firstly, it presents the overall system architecture (section 3.1) paying particular attention to the integration with patient health records, the accessibility at point of care and the modularity in design. To continue, the chapter describes the modules providing decision support (section 3.2), the design strategies (subsection 3.3.1) and the iterative process used to evaluate the usability of the system (subsection 3.3.2). After this, an illustration of the graphical user interface for each of the modules is presented (section 3.3). To conclude, the chapter discusses the main areas in which the CDSS contributes to provide effective care.

1.3.3 Chapter 4: Antimicrobial resistance surveillance

This chapter describes research towards the implementation of a decision support module to facilitate antimicrobial resistance (AMR) surveillance from susceptibility test data. Firstly, the chapter describes the most widespread methods used to measure AMR (section 4.1). Later, it presents a number of strategies to enhance AMR surveillance such as quantification of antimicrobial spectrum of activity (section 4.5), generation of granular and accurate resistance time series (subsection 4.2.2) or estimation of resistance trends through regression analysis (section 4.6). To conclude, the chapter validates the outcomes obtained using the data provided by the Imperial College Healthcare National Health Service Trust in London with those reported by renown health care organizations and independent research articles within the literature.

1.3.4 Chapter 5: Case-based reasoning

This chapter describes research towards the implementation of a decision support module to assist physicians in antimicrobial therapy selection. For such purpose, the chapter describes the case-based reasoning (CBR) methodology (section 5.1). After this, the methodology used to generate the CBR models is explained; from the definition of distance (subsection 5.2.1) and similarity metrics (subsection 5.2.3) to the evaluation of

the performance of the models in the context of information retrieval (subsection 5.2.5) and classification (subsection 5.2.6) problems. In particular, three experiments have been undertaken: an exceedingly simple scenario using the iris flower dataset (section 5.3), an imbalance scenario using the *Escherichia coli* protein location dataset (section 5.4) and an antimicrobial therapy selection trial within to intensive care unit the evaluate the system (section 5.5). To conclude, the findings (section 5.6) and limitations (section 5.6.4) are discussed and the main conclusions summarised (section 5.7).

1.3.5 Chapter 6: Probabilistic inference

This chapter describes research towards the implementation of a decision support module to provide stepwise guidance in infection management. Firstly, the chapter provides a brief introduction to supervised machine learning methods for classification (section 6.1). After this, the implemented methodology to obtain the predictive models is explained; from data assembling (subsection 6.2.2) to model evaluation (subsection 6.2.3). In particular, this methodology has been used to undertake four experiments: prediction of a culture-positive outcome (section 6.4), prediction of the most plausible sites of infection (section 6.5), prediction of culture-positive outcome combining the independent predictions of the most plausible sites (section 6.6) and prediction of the Gram status of bacteria (section 6.7). To conclude, the chapter discusses the findings (section 6.9) and the current limitations (section 6.10).

1.3.6 Chapter 7: Conclusions

This final chapter discusses the benefits of introducing an intelligent CDSS in health care: the convenience of combining various sources of information at the point of care (subsections 7.1.1 and 7.1.2), the importance of providing local AMR surveillance (subsection 7.1.3) and the need to adapt CDSSs to the pathway followed by clinicians in infection management (subsection 7.1.4). To conclude, the chapter summaries the contribution of each decision support module to provide enhanced care (section 7.2) and suggests potential research directions (section 7.3).

BIBLIOGRAPHY

- [1] World Health Organization. "Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016" (2018).
- [2] C Lee Ventola. "The antibiotic resistance crisis: part 1: causes and threats". *Pharmacy and Therapeutics* 40.4 (2015), p. 277.
- [3] C Lee Ventola. "The antibiotic resistance crisis: part 2: management strategies and new agents". *Pharmacy and Therapeutics* 40.5 (2015), p. 344.
- [4] Jim O'Neill. "Antimicrobial resistance: tackling a crisis for the health and wealth of nations". *Review on antimicrobial resistance* (2014), pp. 1–16.
- [5] Alison H Holmes, Luke SP Moore, Arnfinn Sundsfjord, Martin Steinbakk, Sadie Regmi, Abhilasha Karkey, et al. "Understanding the mechanisms and drivers of antimicrobial resistance". *The Lancet* 387.10014 (2016), pp. 176–187.
- [6] Centre for Disease Control and Prevention (CDC) and others. *Antibiotic resistance threats in the United States*. Atlanta: CDC: US Department of Health and Human Services, 2013.
- [7] Laura J Shallcross and Dame Sally C Davies. *Antibiotic overuse: a key driver of antimicrobial resistance*. 2014.
- [8] David L Sackett, William MC Rosenberg, JA Muir Gray, R Brian Haynes, and W Scott Richardson. *Evidence based medicine: what it is and what it isn't*. 1996.
- [9] Margaret A Hamburg and Francis S Collins. "The path to personalized medicine". *New England Journal of Medicine* 363.4 (2010), pp. 301–304.
- [10] Ida Sim, Paul Gorman, Robert A Greenes, R Brian Haynes, Bonnie Kaplan, Harold Lehmann, et al. "Clinical decision support systems for the practice of evidence-based medicine". *Journal of the American Medical Informatics Association* 8.6 (2001), pp. 527–534.
- [11] Yiannis Kyratsis, Raheelah Ahmad, and Alison Holmes. "Technology adoption and implementation in organisations: comparative case studies of 12 English NHS Trusts". *BMJ open* 2.2 (2012), e000872.
- [12] Ton J Cleophas, Aeilko H Zwinderman, and Henny I Cleophas-Allers. *Machine learning in medicine*. Springer, 2013.
- [13] Helen C Davison, Mark EJ Woolhouse, and J Chris Low. "What is antibiotic resistance and how can we measure it?" *Trends in microbiology* 8.12 (2000), pp. 554–559.

-
- [14] Emmanuel E Odjadjare, Etinosa O Igbinosa, Raphael Mordi, Bright Igere, Clara L Igeleke, and Anthony I Okoh. "Prevalence of multiple antibiotics resistant (MAR) *Pseudomonas* species in the final effluents of three municipal wastewater treatment facilities in South Africa". *International journal of environmental research and public health* 9.6 (2012), pp. 2092–2107.
- [15] Krumperman, PH. "Multiple antibiotic resistance indexing of *Escherichia coli* to identify high-risk sources of fecal contamination of foods." *Applied and Environmental Microbiology* 46.1 (1983), pp. 165–170.
- [16] Tandra Mohanta and Sudha Goel. "Prevalence of antibiotic-resistant bacteria in three different aquatic environments over three seasons". *Environmental monitoring and assessment* 186.8 (2014), pp. 5089–5100.
- [17] Charles W Kaspar, Janie L Burgess, Ivor T Knight, and RR Colwell. "Antibiotic resistance indexing of *Escherichia coli* to identify sources of fecal contamination in water". *Canadian journal of microbiology* 36.12 (1990), pp. 891–894.
- [18] Piyush Tripathi, Gopa Banerjee, Shivani Saxena, Mahendra Kumar Gupta, and PW Ramteke. "Antibiotic resistance pattern of *Pseudomonas aeruginosa* isolated from patients of lower respiratory tract infection". *African Journal of Microbiology Research* 5.19 (2011), pp. 2955–2959.
- [19] Mykola Pechenizkiy, Alexey Tsymbal, Seppo Puuronen, Michael Shifrin, and Irina Alexandrova. "Knowledge discovery from microbiology data: many-sided analysis of antibiotic resistance in nosocomial infections". *Biennial Conference on Professional Knowledge Management/Wissensmanagement*. Springer. 2005, pp. 360–372.
- [20] Jessica L Nelson, Barbara L Smith, Jeremy D Jared, and John G Younger. "Prospective trial of real-time electronic surveillance to expedite early care of severe sepsis". *Annals of emergency medicine* 57.5 (2011), pp. 500–504.
- [21] Jungyoun Lee, Seung Sik Hwang, Kyuseok Kim, You Hwan Jo, Jae Hyuk Lee, Joonghee Kim, et al. "Bacteremia prediction model using a common clinical test in patients with community-acquired pneumonia". *The American journal of emergency medicine* 32.7 (2014), pp. 700–704.
- [22] Cara B Litvin, Steven M Ornstein, Andrea M Wessell, Lynne S Nemeth, and Paul J Nietert. "Use of an electronic health record clinical decision support tool to improve antibiotic prescribing for acute respiratory infections: the ABX-TRIP study". *Journal of general internal medicine* 28.6 (2013), pp. 810–816.
- [23] OD William Miller. "A Decision Tree: Proper Antibiotic Selection and Use" (2013).
- [24] Bernhard Heindl, Rainer Schmidt, Gabi Schmid, Mathias Haller, Peter Pfaller, Lothar Gierl, et al. "A case-based consiliarius for therapy recommendation (ICONS): computer based advice for calculated antibiotic therapy in intensive care medicine". *Computer Methods and Programs in Biomedicine* 52.2 (1997), pp. 117–127.
- [25] Nishikant Mishra, Sanja Petrovic, and Santhanam Sundar. "A self-adaptive case-based reasoning system for dose planning in prostate cancer radiotherapy". *Medical physics* 38.12 (2011), pp. 6528–6538.

- [26] Sebastian G Wicha, Martin G Kees, Alexander Solms, Iris K Minichmayr, Alexander Kratzer, and Charlotte Kloft. "TDMx: a novel web-based open-access support tool for optimising antimicrobial dosing regimens in clinical routine". *International journal of antimicrobial agents* 45.4 (2015), pp. 442–444.

BACKGROUND

This chapter confers a general overview on infectious diseases focusing primarily on the current threat of antimicrobial resistance (AMR). In the first instance, a general introduction to infectious diseases is presented (section 2.1). Afterwards, the leading drivers of AMR (subsection 2.2.1), its consequences (subsection 2.2.2) and the potential lines of action (subsection 2.2.3) are explained. Clinical decision support systems (CDSSs) are widely used in health care to support decision-making. Furthermore, they have shown potential to reduce antimicrobial exposure within infection management. For such reason, common underlying infrastructures providing decision support (subsection 2.3.1) and the current state of CDSSs in infection management (subsection 2.3.2) are described. To conclude, the proposed CDSS implemented within this PhD and the individual original contributions included are summarised (subsection 2.4.3).

2.1 INFECTIOUS DISEASES

The human body is colonized by a vast number of microbes collectively referred as the human microbiota. These microbes perform essential functions that contribute to the physiology of the host and are rendered harmless by the human immune system. However, under certain conditions, pathogenic microbes may multiply and invade the host body tissues causing an infection. The illness produced by an infection is called infectious disease and if left untreated (or incorrectly treated) might lead to death. In 2013, more than 9.2 million deaths (about 17% of all deaths) were attributed to infectious diseases [1]. Antimicrobials are used to kill or inhibit the growth of pathogenic microbes and therefore treat infectious diseases. In the last years, the consumption of antimicrobials has increased considerably. Nowadays, with more than 34.8 billion defined daily doses, antimicrobials are among the most commonly prescribed drugs in human medicine [2].

2.1.1 *Infectious microorganisms*

Infectious diseases are commonly classified based on the type of pathogen causing the infection. The most frequent pathogens are viruses, bacteria and fungi which are treated with antivirals, antibacterials (or antibiotics) and antifungals. To begin with, viruses are infectious agents that replicate only within the living cells of other hosts by releasing their genetic material. They typically consist of a nucleic acid molecule and a protein coat. Examples of viral infections are the common cold which affects the upper respiratory tract, viral meningitis which causes an inflammation of the meninges or the

human immunodeficiency virus (HIV) which progressively diminishes the immune system. In contrast, bacteria are unicellular living organisms with complete genetic codes that can replicate with or without a host. Bacteria come in three basic shapes: rod-shaped (bacilli), spherical (cocci) or helical (spirilla). According to the structural differences in their cell walls bacteria might be classified as Gram-positive or Gram-negative. Examples of pathogenic bacteria are *Escherichia coli* and *Staphylococcus aureus*. Some species present high morbidity and mortality with tuberculosis alone causing more than 2 million deaths per year [3]. Last, fungi are multicellular microorganisms that decompose and absorb organic matter causing infections such as the athlete's foot.

2.1.2 Site of infection

Infectious diseases can be also classified according to the organ system affected. These include among other locations the respiratory tract, the urinary tract, the bloodstream, the central nervous system or the skin and soft tissues.

Infection in the urinary system

Urinary tract infections (UTIs) are common infections that occur when pathogenic bacteria enter and multiply in the urinary system. These infections develop more often in women than men and are often classified in lower urinary tract infection which affects the urethra and bladder (cystitis) and upper urinary tract infection which affects ureters and kidneys (pyelonephritis). The latter are less frequent but more severe than the former [4]. Urinary tract infections are common health-care associated infections [5] and therefore a leading cause of health care expenditures. In the USA, approximately 8-10 million people develop a urinary tract infection each year [6].

Infection in the respiratory system

Respiratory tract infections (RTIs) occur when pathogenic bacteria enter and multiply in the lower or upper respiratory tracts. The upper respiratory tract infection affects the ears, nose, sinuses and larynx while the lower respiratory tract infection affects the trachea, bronchial tubes and lungs. The contagiousness of lower respiratory tract infections, which affect especially children and elderly people, are a substantial public health problem with more than 2.74 million deaths attributed worldwide in 2015 [7].

Infection in the bloodstream

The presence of bacteria in the blood, which is normally a sterile environment, is called bacteremia. Bacteria can directly enter the bloodstream through a wound or cut. However, it commonly occurs when underlying infections elsewhere in the body enter the bloodstream. It quickly becomes life-threatening as bacteria, carried through the bloodstream, contaminate other organs away from the original site of infection. The immune system response to bacteremia can cause inflammation throughout the body, originating blood clots that obstruct oxygen delivery to vital organs. The dysregulated host response to an infection might lead to life-threatening organ dysfunction, denoted as sepsis [8]. The epidemiological burden of sepsis is difficult to ascertain, however it is estimated to affect more than 30 million people worldwide each year [9]. The most common underlying conditions causing sepsis are respiratory and urinary infections. If left untreated, it can cause severe sepsis or septic shock. The mortality rates depend on the setting and severity of the disease and though uncertainty still exists, they have been reported to reach up to 30% for sepsis, 50% for severe sepsis and 80% for septic shock [8, 10–12].

2.1.3 Modes of transmission of infectious diseases

The most important routes by which pathogens may spread are: contact, vehicle-borne, air-borne and vector-borne. In healthcare facilities, contact transmission is the most common route [13]. It can occur through direct physical contact, such as during clinical examination of the patient, through indirect contact, such as contaminated medical devices, or by droplet transmission, such as respiratory droplets expelled when a patient coughs or sneezes (e.g. tuberculosis). Furthermore, small droplets and other infectious particles may be carried by dust suspended in the air. This route, called air-borne transmission, can spread infectious diseases over considerable distances without the need of contact. Similarly, vehicle-borne transmission through contaminated objects and substances such as food (e.g. salmonellosis), water (e.g. legionellosis) or blood (e.g. hepatitis) may spread infectious agents. Finally, vector-borne transmission occurs when the insect or animal spreads the infectious agent. While low- and middle-income countries are largely affected by vehicle-borne and vector-borne transmissions, the latter has no significant relevance in health care facilities.

Medical advances have brought lifesaving care to patients in need, yet many of these advances also increase the risk of infection. Despite the development of guidelines for infection control and prevention, the transmission of infectious diseases in healthcare facilities is still a major safety concern [14]. This transmission is often caused by the use of invasive devices such as central lines, urinary catheters or ventilators. The impact of healthcare associated infections results in prolonged hospital stay, additional costs for health care systems and ultimately more deaths which are largely affected by the emerging rates of microbes resistant to previously effective antimicrobials. In Europe, more than 2.6 million cases of healthcare associated infections occur every year [15] which are estimated to cause more than 16 millions extra-days in hospital stay and over 37,000 deaths [16].

2.2 THE BURDEN OF ANTIMICROBIAL RESISTANCE

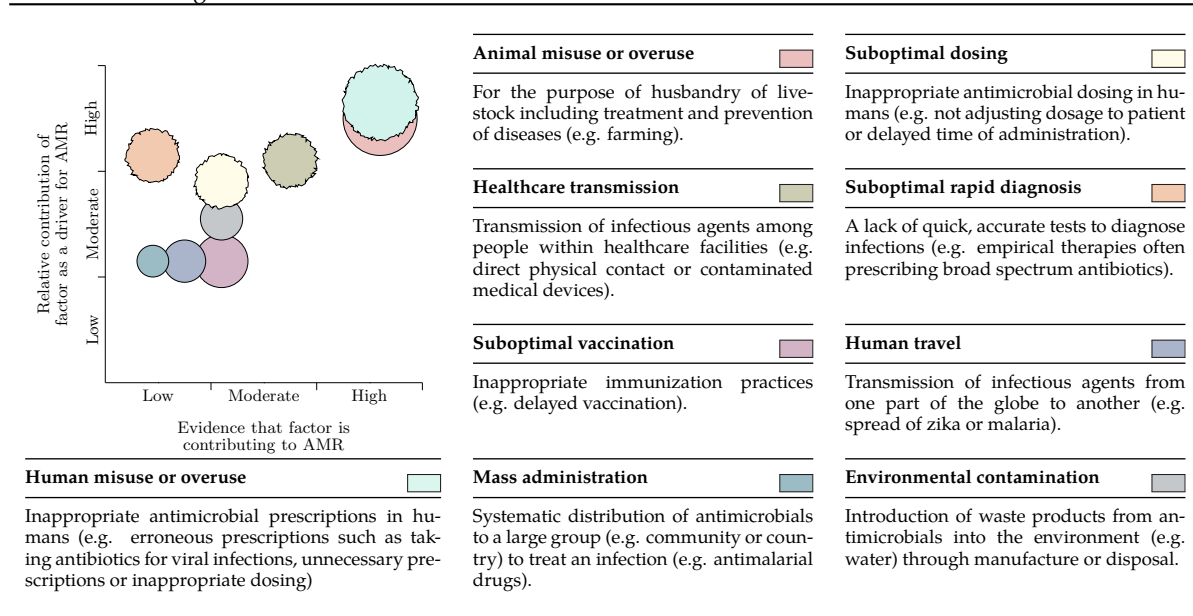
Antimicrobial resistance (AMR) constitutes the ability of microbes to resist antimicrobials which were previously effective. The proportion of resistant microbes has significantly increased in the last years and has become a leading patient health and safety issue, with estimates that it will be responsible for more than 10 million deaths by 2050 [17]. The burden of AMR has received significant attention from high-income countries yet the vast majority of deaths are projected to occur in low- and middle-income countries which generally have to face the lack of financial, material and human resources within their health care systems [17, 18].

2.2.1 What are the main drivers of AMR?

The development of AMR is another manifestation of the Darwinian theory of biological evolution [19]. This theory explains how certain living organisms with variations that increase their ability to compete arise and develop through natural selection. This phenomenon is accelerated by circumstances where selective pressure is exerted. As a result, over the last years, conducted research has focused on identifying the factors contributing to AMR [20–23] and monitoring the disparate levels of resistance around the

world [17]. The alarming outcomes have motivated the emergence of different strategies and guidelines to analyse, present and ultimately combat antimicrobial resistance [24]. A summary of the leading drivers of AMR and their contributions is shown in Table 2.1.

Table 2.1: Leading drivers of antimicrobial resistance.



An info-graphic to show the considered potential contribution of each factor as a driver for antimicrobial resistance including further explanation and examples. The diameter of bubble represents the potential population affected [23, 25]. Irregular circles represent drivers that could be addressed through the implementation of decision support systems in healthcare settings.

A major driver of AMR has been the misuse of antimicrobials in humans [23, 26, 27]. Antimicrobials are among the most commonly prescribed drugs in human medicine, yet up to 50% of all antimicrobials prescribed to people are considered unnecessary [28]. Reasons for the misuse of antimicrobials in humans are complex, yet a number of factors have been described and investigated. At the prescriber level, the majority of antimicrobials are prescribed by individuals who are not experts in infection management and may have limited understanding of antimicrobials and the potential consequences of AMR [23, 29–31]. Furthermore, physicians often prioritise the management of the patient in front of them, paying little regard to the long-term consequences of misusing or overusing antimicrobials [32]. For instance, prescriptions based on clinical educated guess in the absence of complete information are frequent at early stages due to practical shortcomings in rapidly diagnosing infections [33]. These empirical therapies typically rely on the use of broad-spectrum antimicrobials and contribute to the increasing prevalence of multi-resistant microbes [34]. On the other hand, much more widespread is the general over prescribing of antimicrobials by general practitioners. The practice of not completing prescribed courses of antimicrobial treatment, or completing an unnecessary antimicrobial course, may not materially affect the immediate clinical outcome of the patient but can potentially contribute to AMR [23, 25]. To conclude, the non-controlled access to falsified and sub-standard antimicrobials might not only cause harm to patients and fail to treat the diseases for which they are intended but also contribute to the development of AMR [35, 36].

2.2.2 What are the consequences of AMR?

The inability to treat common infectious diseases is an immediate consequence of AMR which leads to various complications and longer hospital stays. The extended hospital stays due to resistant pathogens, the fragile state of the patients and the use of invasive medical devices altogether promote the transmission of infections in health-care settings. At the patient level, this is translated into an increase in morbidity and mortality. At the health care level, it results in an increase of resource utilization and therefore higher costs.

2.2.3 How can we mitigate AMR?

Published work on antimicrobial resistance across many different diseases and disciplines converges on a remarkably consistent set of recommendations for prevention and containment [26, 37] but their implementation has been slow [38]. These recommendations, which are briefly explained below, focus on improvement of diagnosis and prescription practices, reduction of antimicrobial exposure, development of novel antimicrobials, improvement of surveillance and promotion of awareness and education.

Early infection diagnosis at point of care

The development and implementation of rapid and accurate diagnostic tools would alleviate empirical and simultaneous prescriptions [39]. Although this method has been widely advocated [24, 37], its implementation progresses slowly, due partly to technical and financial barriers but also to issues around innovation adoption [40].

Improvement of antimicrobial prescription practices

Although guidelines recommend prudent use, needless prescriptions are seen even in countries with low rates of prescription [41]. However an overall reduction in prescriptions for antimicrobials has been reported in some settings over the past decade with a modest reduction in antimicrobial resistance recorded [39]. The role of educating physicians is crucial in overcoming antimicrobial misuse or overuse and is seen to be effective in primary care [42] and secondary care [43]. Additionally, raising awareness of the fundamentals of antimicrobial use in the general public is equally essential [25].

Sustain useful therapeutic life of antimicrobials

The development of new antimicrobials is substantially difficult and requires large investment to succeed as reflected by the low rate of new antimicrobials released in the last years [44]. Resistance to novel antimicrobials will eventually develop led by natural selection and therefore it is crucial to prolong their useful therapeutic life. Potential approaches to extend the therapeutic life of antimicrobials rely on (i) heterogeneity of antimicrobial agents (ii) combination therapies (iii) use of narrow spectrum antibiotics and (iv) assure adequate drug concentrations [45].

Enhanced mechanisms for surveillance

Surveillance is the cornerstone for assessing the burden of antimicrobial resistance and benefits health care organizations in many different ways: (i) contributes to the evidence base used for formulation of treatment guidelines (ii) can be used to assess the effectiveness and impact of interventions and (iii) has a key role in detecting the emergence and spread of previously uncommon or completely novel types of resistance.

In conclusion, the emergence of antimicrobial resistance is a natural phenomenon, yet it has been accelerated by antimicrobial exposure. To reduce such exposure in health care settings, electronic clinical decision support systems (CDSS) have been devised with the aim of providing physicians with easy and rapid access to information, which is required to make evidence-based therapeutic decisions at the point of care.

2.3 CLINICAL DECISION SUPPORT SYSTEMS

A clinical decision support system (CDSS) is a piece of software designed to help health care professionals make clinical decisions. These systems integrate medical knowledge, patient data and an inference engine to provide case specific advice. In medicine, CDSSs have demonstrated to increase quality of care, enhance health outcomes and reduce human errors by promoting the practice of evidence-based medicine [46].

2.3.1 Overview of decision support engines

The majority of CDSSs are basic systems which provide computerized alerts [47], reminders [48] and electronic clinical guidelines [49]. More elaborated CDSSs may include vital-sign monitoring [50] or hospital-specific statistics [51] through visualization tools. In the last years, machine learning has enabled the creation of decision support systems that learn from past experiences or patterns within the clinical data and are able to stratify patients based on risk [52], support diagnosis [53] or provide therapy advice [54]. This section provides an overview of methodologies for the construction of inferential and predictive CDSSs which have been used in health care settings.

Rule-based systems

In rule-based systems, the expert knowledge is represented in the form of human-crafted rule sets (e.g. IF-THEN rules) that align with the cognitive process followed by humans. These systems require thorough knowledge of the problem being addressed in order to define a comprehensive set of rules. Thus, their coverage is generally low focusing on specific domains. The process of rule generation in complex systems might become challenging and time consuming with large amount of manual effort. Moreover, maintenance tends to be cumbersome, obstructing the inclusion of new developments and emerging initiatives. Rule-based systems have been applied in domains such as antimicrobial therapy selection [54] or diagnosis [55, 56].

Tree-based systems

Decision trees automatically extract rules from data and are popular in clinical domains for their simplicity to understand. The amount of computing time required for large datasets is still reasonable. However, they do not tend to work well if decision boundaries are smooth and are not optimal for uncorrelated variables [57]. As a result of the greedy strategy applied, they also tend to present high variance and are often unstable, tending to over-fit [58]. Tree-based systems have been applied in domains such as dengue fever diagnosis [59] or antimicrobial therapy selection [60].

Case-based reasoning

The case-based reasoning methodology uses previous experience in form of cases to understand and solve new problems [61]. As such, this methodology often requires human interaction to revise those cases which are included into the case base [62]. These systems rely on the definition of a suitable similarity measure to compare cases [63] and the mechanism selected to combine the solutions of the retrieved cases [61]. This approach closely replicates the strategy followed by health care workers during clinical practice. Thus, it has been used to tackle problems on a wide variety of medical domains such as dose planning for cancer radiotherapy [64] or antibiotic therapy selection [65, 66].

Probability-based approaches

Probabilistic methods have emerged due to its capacity to represent and handle uncertainties [67]. The most basic group of probabilistic methods draw conclusions from sample data by emphasizing the frequency or proportion of the data. Bayesian methods are more sophisticated and combine data-driven likelihoods with priors extracted from first principles or intuitions to update the probabilities as more evidence becomes available. One step further, probabilistic graphical models (e.g. Bayesian networks) are probabilistic models for which a graph represents the variables (nodes) and their dependencies (arcs). They offer a natural way of representing uncertainties and can learn the relationship between variables. However, an insufficient understanding of their formal meaning may give rise to modelling flaws. Probabilistic graphical models have been widely exploited in health-care [68, 69] such as diagnosis of ventilator-associated pneumonia [70] or bloodstream infections [53, 71].

2.3.2 CDSSs for infection management

Clinical decision support systems (CDSSs) were first developed to support antimicrobial management in the 1980s. Nowadays, CDSSs assist in areas such as infection disease diagnosis [72] or therapy selection [73]. However, its adoption in infection management is still low [74]. Numerous studies have assessed the effectiveness of CDSSs in antimicrobial management [75, 76] and the appropriateness of such interventions [74]. However, to the best of the author's knowledge, none has reviewed the CDSSs infrastructure (e.g. algorithm or selected attributes) yielding decision support. Towards this objective, an exhaustive literature review to identify the existing approaches to produce curated decision-support was conducted. A total of 68 studies with 53 different CDSSs were considered. Exclusively CDSSs implemented in health-care settings which provided decision support during the patient evaluation process were included. For the sake of clarity, the main properties of the CDSSs are summarised in Table 2.2. The full evaluation is provided as supplementary material (see Appendix B.1).

The majority of CDSS in the literature are rule-based systems (44/68; 64%). These are largely focused on replicating paper-based guidelines in electronic format to ensure adherence (24/44; 57%). Rule-based systems outnumbered those incorporating more advanced machine learning methods such as tree-based (12/68; 18%), case-based reasoning (7/68; 10%) or Bayesian networks (1/68; 1%). The main driver was clinical confidence in the system to increase CDSS adoption [77]. Antimicrobial therapy selection was the primary outcome (47/68; 69%) with few addressing surveillance (7/68; 10%), infection diagnosis (5/68; 7%) or antimicrobial dosing (2/68; 3%). Furthermore, studies providing more than one type of decision support were low (2/68; 3%). The conditions

targeted were respiratory tract infections (14/68; 21%), bloodstream infections (7/68; 10%) and urinary tract infections (3/68; 4%). A small number of studies (6/68; 9%) tackled more than one domain [72, 73]. Several platforms for delivering decision support were reported, including EHR add-on modules (31/68; 45%), standalone software (11/68; 16%), web-based (8/68; 12%) or personal digital assistants (3/68; 4%).

The majority of CDSSs implemented in health care settings are at secondary care (53/68; 78%). In general, the selected targets are the intensive care unit and the emergency department with most systems relying on either human-crafted rules or tree-based methods. The use of more advanced machine learning methods to provide decision support is significant in research (see appendices B.1 and B.2). However, these are rarely translated into practical applications (12/68; 18%). Note that these research studies generally assume no data constraints during the design and development of the models and therefore use a large number of variables (often more than 50) which are extremely difficult to collect in clinical practice. In addition, the data used for validation is often scant (from 100 to 500 cases with few exceptions [72]). As a generic rule, the amount of data required to obtain reliable results grows exponentially with the dimensionality.

Despite the benefits that CDSSs may bring to clinical practice their adoption remains low [40]. Among the technical factors limiting CDSS adoption are accessibility at point of care, ongoing technical problems or the lack of integration within existing systems [78]. Moreover, integration within clinical workflow, ease of use and quality of the information have been reported to impact CDSS uptakes [77, 78]. For instance, generic CDSSs that did not account for local constraints were often rejected [78]. On the other hand, a number of personal and social factors have been identified as barriers to adoption. These factors include concerns about liability and privacy, feelings that systems do not respect the autonomy of the physician or the possible effects that these systems might have on patient-doctor interactions [79]. A clear facilitator was the inclusion of end users into the development process and the support of senior doctors to encourage use and provide users training for the system [78]. To conclude, user expectations need to be taken into consideration and should be evaluated in the early phases and throughout the development of these systems to increase adoption [80, 81]. For instance, moving away from a black-box process to a more transparent method that describes the physicians how the system is making the decision has been suggested as a potential facilitator [82].

To summarise, the most unexpected outcomes and limitations identified from the previous literature review are: (i) the majority of CDSSs employ human-crafted rules and only a few implement more advanced machine learning techniques to extract patterns from data (ii) a small number of CDSSs tackle infection diagnosis (iii) a small number of CDSSs address patient involvement and education (iv) the retrospective and/or prospective validation is often insufficient and (v) the solutions are utterly complex with high loads of input data required.

2.3.3 *Methods for infectious disease diagnosis*

Based upon the previously identified weaknesses, a review of the literature to evaluate existing methods to diagnose infectious diseases using machine learning was carried out. A total of 42 studies were considered. These studies stand as research articles without existing implementation in health care settings. The properties of these studies are summarised in Table 2.3. The full evaluation is provided as supplementary material (see Appendix B.2).

The majority of conducted research in infectious disease diagnosis focuses on bloodstream infection (20/42; 49%), followed by surgical site infection (8/42; 19%), urinary tract infection (6/42; 14%) and respiratory tract infection (6/42; 14%). The most common approach to diagnose bacteremia relies on regression analysis combined with rule-based methods. These studies consider a reduced set of clinical variables (less than 10) which are predominantly bio-chemical markers such as c-reactive protein or procalcitonin. A mean area under the ROC of 0.75 was reported with several studies failing to describe sensitivity and specificity [53, 56, 83]. When described, these were disparate with a geometric mean of 0.68. Another popular approach relies on tree-based methods with a mean area under the ROC of 0.88 0.75 with a geometric mean of 0.71. Since 2015, the contribution of machine learning has emerged with support vector machines among the most popular methods. These achieved a mean area under the ROC of 0.81 and a geometric mean of 0.72.

In general, the best results for infection diagnosis were obtained when time series data was available (5/8; 63%). Note that the inclusion basis within data collection was always very restrictive. For instance, most patient had to fulfil conditions such as no antimicrobial therapy dispensed, suspicion of infection or other conditions such as the systemic inflammatory response syndrome (SIRS) or sequential organ failure assessment (SOFA). A large number of clinical variables was often required (from 30 to 200) including symptoms, vital signs, biochemical markers, microbiology outcome or doctors notes. The algorithm validation was performed in most cases retrospectively and the issue of class imbalance within the data, recurrent in these type of studies, was never addressed.

Table 2.2: Literature review summary: comparison of clinical decision support systems for infection management.

System setting		System domain		Type of decision support		Platform		System infrastructure	
place or location in which the system was trialled		syndrome, condition or infectious disease targeted by the system		duty or task for which the system provides assistance to clinicians		application or frame in which suggestions are displayed		underlying technique analysing the data to generate suggestions	
Primary care	15 (22%)	Bloodstream infection	7 (10%)	Antimicrobial selection	47 (69%)	EHR add-on	21 (31%)	Rule-based	44 (64%)
Secondary care	53 (78%)	Respiratory tract infection	14 (21%)	Infection diagnosis	5 (5%)	On PDA device	3 (4%)	Tree-based	12 (18%)
		Urinary tract infection	3 (4%)	Knowledge sharing	1 (2%)	Web-based application	8 (12%)	Case-based reasoning	7 (10%)
		Surgical site infection	2 (3%)	Antimicrobial dosing	2 (3%)	Standalone software	11 (16%)	Bayesian networks	1 (1%)
		Other	9 (13%)	Surveillance	7 (10%)	N/A	25 (25%)	PKPD modelling	1 (1%)
		Multiple	6 (9%)	Multiple	2 (3%)			Other	2 (3%)
		N/A	27 (40%)	Other	4 (6%)			N/A	1 (1%)

Keys: EHR=electronic health records; PDA=Personal digital assistant; PKPD=Pharmacokinetics/Pharmacodynamics; n/a=not applicable.

Table 2.3: Literature review summary: comparison of methodologies for infectious disease diagnosis.

Methodology	Bloodstream	Respiratory tract	Urinary tract	Surgical site	Comments I		Comments II	
the underlying technique analysing the data to provide the diagnosis outcome	domains such as sepsis and septic shock	domains such as infection by pneumonia	domains such as UTI infection	domains such as post-operative infections	this section describes the type of diagnosis and the methodology used		this section describes the data, the type of variables and the number of variables within the studies	
Regression analysis					Type of diagnosis		Type of variables [†]	
AUCROC	0.75 (±0.08)	0.90 (±0.01) [‡]	0.83 (±0.12)	0.81 (±0.09)	Bacteremia	20 (49%)	Symptoms	7 (10%)
SENS	0.68 (±0.17)	0.60 (±0.25)	0.65 (±0.15)	0.38 (±0.23)	Urinary tract infection	6 (14%)	Vital signs	11 (16%)
SPEC	0.72 (±0.20)	0.75 (±0.14)	0.78 (±0.22)	0.94 (±0.05)	Respiratory tract infection	6 (14%)	Demographics	14 (20%)
GMEAN	0.68 (±0.07)	0.67 (±0.19)	0.70 (±0.11)	0.60 (±0.27)	Surgical site infection	8 (19%)	Biochemical markers	20 (29%)
					Culture positivity	1 (2%)	Microbiology	6 (9%)
					Gram status	1 (2%)	Others	12 (17%)
Tree-based methods					Methodology		Number of variables	
AUCROC	0.88 (±0.08)	n/a	0.89 (±0.02)	n/a	Score	3 (7%)	Score	11–17 (M)
SENS	0.66 (±0.25)	n/a	0.57 (±0.03)	n/a	Regression analysis	18 (41%)	Regression analysis	1–11 (L)
SPEC	0.81 (±0.19)	n/a	0.95 (±0.01)	n/a	Gaussian approach	1 (2%)	Gaussian approach	10–12 (M)
GMEAN	0.71 (±0.07)	n/a	0.74 (±0.02)	n/a	Tree-based methods	8 (20%)	Tree-based methods	10–52 (H)
Support Vector Machines					Support vector machine	4 (10%)	Support vector machine	10–52 (M)
AUCROC	0.81 (±0.09)	n/a	0.85 (±0.04)	n/a	Artificial neural network	4 (10%)	ANN	30–50 (H)
SENS	0.60 (±0.03)	n/a	0.50 (±0.00)	n/a	Other	4 (10%)		
SPEC	0.87 (±0.13)	n/a	0.97 (±0.01)	n/a			Data	
GMEAN	0.72 (±0.07)	n/a	0.69 (±0.00)	n/a			Imbalanced	41 (100%)
Artificial Neural Networks							Number of cases*	
AUCROC	n/a	n/a	0.88 (±0.01) [‡]	0.85 (±0.06) [‡]			<100	1 (2%)
SENS	n/a	n/a	0.54 (±0.01) [‡]	0.64 (±0.04) [‡]			<1000	15 (20%)
SPEC	n/a	n/a	0.95 (±0.00) [‡]	0.92 (±0.04) [‡]			<10000	17 (10%)
GMEAN	n/a	n/a	0.72 (±0.01) [‡]	0.77 (±0.04) [‡]			<50000	7 (10%)
							>50000	2 (10%)

Notation: mean score (±std); *= classes above are excluded; †=several type of variables per study; ‡=low number of studies to compute the values.

Keys: AUCROC=area under the receiver operating characteristic curve; SENS=sensitivity; SPEC=specificity; GMEAN=geometric mean; ANN=artificial neural network; L=low; M=medium; H=high;

2.4 TOWARDS COMPREHENSIVE DECISION SUPPORT

A wide variety of clinical domains benefit from data-driven decision support systems, yet their introduction within infection management has been much slower [74]. After reviewing the use of CDSSs in infection management (see Appendix B.1), numerous factors limiting their adoption were identified. Among these factors, the lack of integration with electronic health records [84], the need of accessibility at point of care [85] and the complexity of the underlying systems [86, 87] deserve special attention. In addition, the review conducted on infectious diseases diagnosis (see Appendix B.2) revealed the enormous potential of using machine learning to extract patterns from clinical data. Furthermore, these methods could escalate the level of assistance provided to clinicians through the different steps of the infection management pathway (e.g. diagnosis, therapy selection or dosing) [88]. These findings have motivated the development of a comprehensive and methodical CDSS to facilitate and improve infection management at point of care.

2.4.1 The five ‘rights’ of a CDSS

The concept of CDSS has been important in health care for many years. Traditionally, CDSSs had a content-centric role and has expanded to include a variety of sources and services. While the ubiquity of information technology in health care has promoted the development of advanced CDSSs, the following ‘rights’ (see Table 2.4) continue to apply to provide effective decision support [89]. These five rights should be used as a framework when implementing a CDSS.

Table 2.4: The five rights of a CDSS: brief description and examples.

The right information	To the right person	In the right format
Evidence-based information suitable to guide action according to the circumstance (e.g. use symptoms, laboratory tests and previous experience to determine further testing, diagnose the infectious disease or guide antimicrobial selection).	Consider members of the care team, including clinicians, patients and their caretakers to present data accordingly (e.g. facilitate or enhance patient monitoring to nurses or provide local resistance patterns to infection specialists).	Understand how the data should be presented (e.g. a concise infection risk score or probability, a comprehensive summary of the symptoms, a detailed temporal evolution of the biochemical markers or links to external information).
Through the right channel	At the right time	
Related to the system used to provide the support (e.g. a clinical information system on a computer or a mobile device, a pop up alert or a simple document that can be printed or attached in an email).	Align with the infection management workflow to provide support at time of decision (e.g. specimen collection, diagnosis or antimicrobial selection at point of care or notification for de-escalation of therapy).	

2.4.2 The desired characteristics of a CDSS

CDSSs represent a vital component of the healthcare enterprise and have evolved from paper-based and content-driven to digital systems that offer all medical stakeholders access to analytic tools for critical decision making. These can be classified according to their topic in clinical guidelines, evidence-based medicine, analytic tools or point of care solutions. To increase the potential of CDSS in infection management a number of desired attributes (see Table 2.5) and functionality (see Table 2.6) have been identified.

Table 2.5: Desired attributes of a CDSS for infection management.

Accessibility at point of care	Integration with EHR	Modularity in design
to provide decision support timely and facilitate interpersonal communication (e.g. during patient examination)	to facilitate automated collection of patient related data (e.g. pathology laboratory tests or microbiology data)	to reduce complexity and facilitate maintenance (e.g. incorporate updated guidelines, integrate novel diagnostic tests or upgrade algorithms)
Availability of data	Align with decision pathway	Comprehensibility
to facilitate the implementation in health care systems, a reduced number of parameters which are already available or easy to collect must be selected	to provide decision support through the infection management pathway (e.g. request specimen collection, confirm diagnosis, select antimicrobial therapy and dosing or assess the likelihood of success)	to facilitate innovation adoption the system needs to be easy to understand and the interface readable and clear

Table 2.6: Desired functionality of a CDSS for infection management.

Facilitate vital sign collection	Patient monitoring	Patient engagement
to provide a homogeneous set of vital signs for comparison that can be further exploited through machine learning algorithms to enhance decision support	to observe the disease, medical condition and other vital parameters over time to prevent or early detect adverse events	to improve their understanding of antimicrobial use and shape future patient behaviour
Review of past similar cases	Therapeutic drug monitoring	Local AMR surveillance
to provide the necessary information to improve antimicrobial therapy selection matching the natural reasoning and cognitive process followed by clinicians	to improve patient care by adjusting the antimicrobial dose to increase effectiveness while reducing the secondary toxic effects	to strengthen knowledge for action in support of strategies to improve existing guidelines (e.g. antimicrobial spectrum of activity or resistance rates and trends)

For health care providers, electronic health records (EHR) systems are the preferred method of data implementation and has quickly become the backbone of healthcare organizations [84, 90, 91]. As such, CDSSs integration into EHR is no longer a luxury and has become a highly desired feature. In addition, the adoption of EHR promotes the development of technologies that are accessible at the point of care to minimize the time spent on documentation, facilitate communication between health care workers and provide decision support that fits in the workflow. In general, CDSSs in infection management focus predominantly on antimicrobial therapy selection. However, decision support is much needed on other areas of the infection management pathway. At patient admission, it is crucial to smooth vital sign recording, guide further specimen collection and diagnose the infectious disease. Furthermore, the medical history of the patient and previous similar cases can provide clinicians with the evidence needed for appropriate empirical antimicrobial therapy selection. Once the treatment is applied, patient monitoring is essential to evaluate the response of the disease and to identify adverse drug reactions. Integrating all the previously mentioned functionality into a single CDSS could lead to an increase in the perceived complexity of the system which might obstruct clinical acceptance. Thus, to provide flexibility in design and reduce the perceived complexity of the system, a partition of the CDSS into scalable and reusable modules with a well defined function should be considered. Modularity in design also facilitates maintenance and further integration of novel algorithms or diagnostic devices.

2.4.3 The proposed solution

The aim of this investigation is to understand the needs within infection management and use such knowledge to develop a comprehensive decision support system. As a result, this research has produced Enhanced, Personalized and Integrated Care for Infection Management at Point of Care (EPiC IMPOC) which is a data-driven clinical decision support system that aims to overcome many of the challenges within infection management. As such, EPiC IMPOC has been designed to include the identified desired attributes (see Table 2.5) and functionality (see Table 2.6). In addition, it is in compliance with the five rights. The system incorporates two decision support engines: case-based reasoning and probabilistic inference. The former facilitates data collection, patient monitoring and inspection of previous similar cases to assist in antimicrobial therapy selection. The latter produces stepwise guidance to customize medical decisions to the individual patient to assist with infection diagnosis. Moreover, hospital-specific antimicrobial resistance data is provided to enhance awareness and education among healthcare workers. To conclude, EPiC IMPOC has been integrated with the electronic health records (EHR) and it is accessible at the point of care to improve consistency of collecting vital signs, support interpersonal communication and facilitate interactions between clinicians, infection specialists and patients.

The process followed to design and implement the system and the modules included within EPiC IMPOC are thoroughly explained in subsequent chapters. For the sake of clarity, a brief description of the main challenges addressed and the proposed implementation are briefly described below.

- *Provision of decision support at the point of care (POC)*: a web-based client application has been developed to facilitate access to EPiC IMPOC from any device at POC. The web-based application renders a consistent interface across different screen sizes and can be used in desktop computers or hand-held devices.
- *Integration with electronic health care records (EHR)*: the system has been integrated with the electronic health records to automatically create profiles for the admitted patients. In addition, it has been interfaced with several National Health Services (NHS) databases to display laboratory results at point of care.
- *Antimicrobial resistance (AMR) surveillance*: to promote awareness and understanding of resistance patterns, hospital-specific AMR statistics such as resistance rates, resistance trends or antimicrobial spectrum of activity are computed automatically from susceptibility test data. Note that these patterns might vary geographically.
- *Evaluation of previous similar cases for therapeutic advice*: case-based reasoning (CBR) has been selected to facilitate homogeneous collection of vital signs, patient monitoring and knowledge transfer among healthcare workers. More importantly, it provides therapeutic advice for antimicrobial selection based on previous experiences.
- *Assessing the likelihood of infection*: a number of supervised machine learning methods have been evaluated to provide for a given patient the likelihood of having an infection. This information is extracted from a set of commonly requested biochemical markers and therefore is presented before the initiation of the antimicrobial therapy.

- *Assessing the site of infection:* a number of supervised machine learning methods have been evaluated to provide for a given patient the most plausible sites of infection. This information guides health care workers planning further investigations such as which microbiology cultures should be requested (e.g blood, sputum or urine).
- *The principle of modularity in design:* to facilitate adoption and maintenance, each of the tasks for which assistance is provided has an independent inference engine. Therefore, the algorithm and explanatory variables can be easily modified to improve performance or just adapt models to use those attributes (e.g. vital signs or biochemical markers) provided within an specific healthcare centre.

As discussed in the literature review, the majority of CDSS for infection management focus on antimicrobial therapy selection. While this is an important decision to be made by physicians, other fundamental questions need to be addressed before antimicrobials are prescribed. As such, a comprehensive CDSS for infection management needs to align with the process followed by clinicians, facilitate communication between health care workers from nurses to infectious specialists and integrate support for the various types of infection to improve the provision of efficient care.

BIBLIOGRAPHY

- [1] II Abubakar, T Tillmann, and A Banerjee. "Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013". *Lancet* 385.9963 (2015), pp. 117–171.
- [2] Eili Y Klein, Thomas P Van Boeckel, Elena M Martinez, Suraj Pant, Sumanth Gandra, Simon A Levin, et al. "Global increase and geographic convergence in antibiotic consumption between 2000 and 2015". *Proceedings of the National Academy of Sciences* (2018), p. 201717295.
- [3] World Health Organization. *Global tuberculosis report 2017*. World Health Organization, 2017.
- [4] Ana L Flores-Mireles, Jennifer N Walker, Michael Caparon, and Scott J Hultgren. "Urinary tract infections: epidemiology, mechanisms of infection and treatment options". *Nature reviews microbiology* 13.5 (2015), p. 269.
- [5] European Centre for Disease Prevention and Control. *Annual epidemiological report 2014. Antimicrobial resistance and healthcare-associated infections*. 2015.
- [6] Vitaly Smelov, Kurt Naber, and Truls E Bjerklund Johansen. "Improved classification of urinary tract infection: future considerations". *European Urology Supplements* 15.4 (2016), pp. 71–80.
- [7] Christopher Troeger, Mohammad Forouzanfar, Puja C Rao, Ibrahim Khalil, Alexandria Brown, Scott Swartz, et al. "Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015". *The Lancet Infectious Diseases* 17.11 (2017), pp. 1133–1161.
- [8] Mervyn Singer, Clifford S Deutschman, Christopher Warren Seymour, Manu Shankar Hari, Djillali Annane, Michael Bauer, et al. "The third international consensus definitions for sepsis and septic shock (Sepsis-3)". *Jama* 315.8 (2016), pp. 801–810.
- [9] Carolin Fleischmann, André Scherag, Neill KJ Adhikari, Christiane S Hartog, Thomas Tsaganos, Peter Schlattmann, et al. "Assessment of global incidence and mortality of hospital-treated sepsis. Current estimates and limitations". *American journal of respiratory and critical care medicine* 193.3 (2016), pp. 259–272.
- [10] Issrah Jawad, Ivana Lukšić, and Snorri Bjorn Rafnsson. "Assessing available information on the burden of sepsis: global estimates of incidence, prevalence and mortality". *Journal of global health* 2.1 (2012).
- [11] Kevin B Laupland. "Incidence of bloodstream infection: a review of population-based studies". *Clinical microbiology and infection* 19.6 (2013), pp. 492–500.

- [12] Elizabeth K Stevenson, Amanda R Rubenstein, Gregory T Radin, Renda Soylemez Wiener, and Allan J Walkey. "Two decades of mortality trends among patients with severe sepsis: a comparative meta-analysis". *Critical care medicine* 42.3 (2014), p. 625.
- [13] Amy S Collins. "Preventing health care-associated infections" (2008).
- [14] Yatin Mehta, Abhinav Gupta, Subhash Todi, SN Myatra, DP Samaddar, Vijaya Patil, et al. "Guidelines for prevention of hospital acquired infections". *Indian journal of critical care medicine: peer-reviewed, official publication of Indian Society of Critical Care Medicine* 18.3 (2014), p. 149.
- [15] Benedetta Allegranzi, Claire Kilpatrick, Julie Storr, Edward Kelley, Benjamin J Park, Liam Donaldson, et al. "Global infection prevention and control priorities 2018–22: a call for action". *The Lancet Global Health* 5.12 (2017), e1178–e1180.
- [16] World Health Organization et al. "Report on the burden of endemic health care-associated infection worldwide" (2011).
- [17] Jim O'Neill. "Antimicrobial resistance: tackling a crisis for the health and wealth of nations". *Review on antimicrobial resistance* (2014), pp. 1–16.
- [18] World Health Organization. "Global action plan on antimicrobial resistance". *Journal of global health* (2015).
- [19] Alex Wong. "Epistasis and the evolution of antimicrobial resistance". *Frontiers in microbiology* 8 (2017), p. 246.
- [20] Peter Collignon, Prema-chandra Athukorala, Sanjaya Senanayake, and Fahad Khan. "Antimicrobial resistance: the major contribution of poor governance and corruption to this growing problem". *PLoS One* 10.3 (2015), e0116746.
- [21] Fred C Tenover. "Mechanisms of antimicrobial resistance in bacteria". *The American journal of medicine* 119.6 (2006), S3–S10.
- [22] Louis B Rice. "Antimicrobial resistance in gram-positive bacteria". *American journal of infection control* 34.5 (2006), S11–S19.
- [23] Alison H Holmes, Luke SP Moore, Arnfinn Sundsfjord, Martin Steinbakk, Sadie Regmi, Abhilasha Karkey, et al. "Understanding the mechanisms and drivers of antimicrobial resistance". *The Lancet* 387.10014 (2016), pp. 176–187.
- [24] Department of Health. *UK five year antimicrobial resistance strategy: 2013 to 2018*. London; 2013.
- [25] Enrique Castro-Sánchez, Luke SP Moore, Fran Husson, and Alison H Holmes. "What are the factors driving antimicrobial resistance? Perspectives from a public event in London, England". *BMC infectious diseases* 16.1 (2016), p. 465.
- [26] Ramanan Laxminarayan, Adriano Duse, Chand Wattal, Anita KM Zaidi, Heiman FL Wertheim, Nithima Sumpradit, et al. "Antibiotic resistance—the need for global solutions". *The Lancet infectious diseases* 13.12 (2013), pp. 1057–1098.
- [27] World Health Organization et al. *The evolving threat of antimicrobial resistance: options for action*. Geneva: World Health Organization, 2012.
- [28] Centers for Disease Control, Prevention (CDC), et al. *Antibiotic resistance threats in the United States, 2013*. Atlanta: CDC; 2013. 2014.

- [29] C Pulcini, F Williams, N Molinari, P Davey, and D Nathwani. "Junior doctors' knowledge and perceptions of antibiotic resistance and prescribing: a survey in France and Scotland". *Clinical microbiology and infection* 17.1 (2011), pp. 80–87.
- [30] António Teixeira Rodrigues, Fátima Roque, Amílcar Falcao, Adolfo Figueiras, and Maria Teresa Herdeiro. "Understanding physician antibiotic prescribing behaviour: a systematic review of qualitative studies". *International journal of antimicrobial agents* 41.3 (2013), pp. 203–212.
- [31] Michael P Doyle, Guy H Loneragan, H Morgan Scott, and Randall S Singer. "Antimicrobial resistance: challenges and perspectives". *Comprehensive Reviews in Food Science and Food Safety* 12.2 (2013), pp. 234–248.
- [32] Jasper Littmann and AM Viens. "The ethical significance of antimicrobial resistance". *Public health ethics* 8.3 (2015), pp. 209–224.
- [33] Almudena Burillo and Emilio Bouza. "Use of rapid diagnostic techniques in ICU patients with infections". *BMC infectious diseases* 14.1 (2014), p. 593.
- [34] Carl Llor and Lars Bjerrum. "Antimicrobial resistance: risk associated with antibiotic overuse and initiatives to reduce the problem". *Therapeutic advances in drug safety* 5.6 (2014), pp. 229–241.
- [35] Emily Leung, Diana E Weil, Mario Raviglione, and Hiroki Nakatani. "The WHO policy package to combat antimicrobial resistance". *Bulletin of the World Health Organization* 89 (2011), pp. 390–392.
- [36] James A Ayukekbong, Michel Ntemgwa, and Andrew N Atabe. "The threat of antimicrobial resistance in developing countries: causes and control strategies". *Antimicrobial Resistance & Infection Control* 6.1 (2017), p. 47.
- [37] European Centre for Disease Prevention and Control. *Surveillance of antimicrobial resistance in Europe 2016. Annual report of the European Antimicrobial Resistance Surveillance Network (EARS-Net)*. Stockholm: ECDC; 2017.
- [38] Pia Abel zur Wiesch, Roger Kouyos, Jan Engelstädter, Roland R Regoes, and Sebastian Bonhoeffer. "Population biological principles of drug-resistance evolution in infectious diseases". *The Lancet infectious diseases* 11.3 (2011), pp. 236–247.
- [39] David M Livermore, Russell Hope, Rosy Reynolds, Ruth Blackburn, Alan P Johnson, and Neil Woodford. "Declining cephalosporin and fluoroquinolone non susceptibility among bloodstream Enterobacteriaceae from the UK: links to prescribing change?" *Journal of Antimicrobial Chemotherapy* 68.11 (2013), pp. 2667–2674.
- [40] Yiannis Kyratsis, Raheelah Ahmad, and Alison Holmes. "Technology adoption and implementation in organisations: comparative case studies of 12 English NHS Trusts". *BMJ open* 2.2 (2012), e000872.
- [41] SR Arnold and SE Straus. "Interventions to improve antibiotic prescribing practices in ambulatory care". *Evidence-Based Child Health: A Cochrane Review Journal* 1.2 (2006), pp. 623–690.
- [42] Paul Little, Beth Stuart, Nick Francis, Elaine Douglas, Sarah Tonkin-Crine, Sibyl Anthierens, et al. "Effects of internet-based training on antibiotic prescribing rates for acute respiratory-tract infections: a multinational, cluster, randomised, factorial, controlled trial". *The Lancet* 382.9899 (2013), pp. 1175–1182.

-
- [43] Peter Davey, Erwin Brown, Esmita Charani, Lynda Fenelon, Ian M Gould, Alison Holmes, et al. "Interventions to improve antibiotic prescribing practices for hospital ins". *Cochrane Database Syst Rev* 4.4 (2013).
- [44] Anthony RM Coates, Gerry Halls, and Yanmin Hu. "Novel classes of antibiotics or more of the same?" *British journal of pharmacology* 163.1 (2011), pp. 184–194.
- [45] Ignasi Roca, Murat Akova, F Baquero, J Carlet, M Cavaleri, S Coenen, et al. "The global threat of antimicrobial resistance: science for intervention". *New microbes and new infections* 6 (2015), pp. 22–29.
- [46] Ida Sim, Paul Gorman, Robert A Greenes, R Brian Haynes, Bonnie Kaplan, Harold Lehmann, et al. "Clinical decision support systems for the practice of evidence-based medicine". *Journal of the American Medical Informatics Association* 8.6 (2001), pp. 527–534.
- [47] Steven L Bernstein, David Whitaker, Jonathan Winograd, and John A Brennan. "An Electronic Chart Prompt to Decrease Proprietary Antibiotic Prescription to Self-Pay Patients". *Academic emergency medicine* 12.3 (2005), pp. 225–231.
- [48] Bala G Nair, Shu-Fang Newman, Gene N Peterson, and Howard A Schwid. "Automated electronic reminders to improve redosing of antibiotics during surgical cases: comparison of two approaches". *Surgical infections* 12.1 (2011), pp. 57–63.
- [49] Elisa Demonchy, Jean-Charles Dufour, Jean Gaudart, Emmanuel Cervetti, Pierre Michelet, Nicolas Poussard, et al. "Impact of a computerized decision support system on compliance with guidelines on antibiotics prescribed for urinary tract infections in emergency departments: a multicentre prospective before-and-after controlled interventional study". *Journal of Antimicrobial Chemotherapy* 69.10 (2014), pp. 2857–2863.
- [50] Jessina C McGregor, Elizabeth Weekes, Graeme N Forrest, Harold C Standiford, Eli N Perencevich, Jon P Furuno, et al. "Impact of a computerized clinical decision support system on reducing inappropriate antimicrobial use: a randomized controlled trial". *Journal of the American Medical Informatics Association* 13.4 (2006), pp. 378–384.
- [51] Signe Flottorp, Andrew D Oxman, Kari Håvelsrud, Shaun Treweek, and Jeph Herin. "Cluster randomised controlled trial of tailored interventions to improve the management of urinary tract infections in women and sore throat". *Bmj* 325.7360 (2002), p. 367.
- [52] Charles J Mullett, John G Thomas, Connie L Smith, Arif R Sarwari, and Rashida A Khakoo. "Computerized antimicrobial decision support: an offline evaluation of a database-driven empiric antimicrobial guidance program in hospitalized patients with a bloodstream infection". *International journal of medical informatics* 73.5 (2004), pp. 455–460.
- [53] Mical Paul, Steen Andreassen, Anders D Nielsen, Evelina Tacconelli, Nadja Al-manasreh, Abigail Fraser, et al. "Prediction of bacteremia using TREAT, a computerized decision-support system". *Clinical Infectious Diseases* 42.9 (2006), pp. 1274–1282.

- [54] Cara B Litvin, Steven M Ornstein, Andrea M Wessell, Lynne S Nemeth, and Paul J Nietert. "Use of an electronic health record clinical decision support tool to improve antibiotic prescribing for acute respiratory infections: the ABX-TRIP study". *Journal of general internal medicine* 28.6 (2013), pp. 810–816.
- [55] Jessica L Nelson, Barbara L Smith, Jeremy D Jared, and John G Younger. "Prospective trial of real-time electronic surveillance to expedite early care of severe sepsis". *Annals of emergency medicine* 57.5 (2011), pp. 500–504.
- [56] Jungyoup Lee, Seung Sik Hwang, Kyuseok Kim, You Hwan Jo, Jae Hyuk Lee, Joonghee Kim, et al. "Bacteremia prediction model using a common clinical test in patients with community-acquired pneumonia". *The American journal of emergency medicine* 32.7 (2014), pp. 700–704.
- [57] J. Ross Quinlan. "Induction of decision trees". *Machine Learning* 1.1 (1986), pp. 81–106.
- [58] Bhumika Gupta, Aditya Rawat, Akshay Jain, Arpit Arora, and Naresh Dhama. "Analysis of Various Decision Tree Algorithms for Classification in Data Mining". *International Journal of Computer Applications* 163.8 (2017).
- [59] Lukas Tanner, Mark Schreiber, Jenny GH Low, Adrian Ong, Thomas Tolfvenstam, Yee Ling Lai, et al. "Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness". *PLoS Negl Trop Dis* 2.3 (2008), e196.
- [60] OD William Miller. "A Decision Tree: Proper Antibiotic Selection and Use" (2013).
- [61] Agnar Aamodt and Enric Plaza. "Case-based reasoning: Foundational issues, methodological variations, and system approaches". *AI communications* 7.1 (1994), pp. 39–59.
- [62] Bernard Hernandez, Pau Herrero, Timothy M. Rawson, Luke S. P. Moore, Esmita Charani, Alison H. Holmes, et al. "Data-driven Web-based Intelligent Decision Support System for Infection Management at Point-Of-Care: Case-Based Reasoning Benefits and Limitations". 5 (2017), pp. 119–127.
- [63] T Warren Liao, Zhiming Zhang, and Claude R Mount. "Similarity measures for retrieval in case-based reasoning systems". *Applied Artificial Intelligence* 12.4 (1998), pp. 267–288.
- [64] Lothar Gierl, Dagmar Steffen, Dusan Ihracky, and Rainer Schmidt. "Methods, architecture, evaluation and usability of a case-based antibiotics advisor". *Computer methods and programs in biomedicine* 72.2 (2003), pp. 139–154.
- [65] Bernhard Heindl, Rainer Schmidt, Gabi Schmid, Mathias Haller, Peter Pfaller, Lothar Gierl, et al. "A case-based consiliarius for therapy recommendation (ICONS): computer based advice for calculated antibiotic therapy in intensive care medicine". *Computer Methods and Programs in Biomedicine* 52.2 (1997), pp. 117–127.
- [66] Nishikant Mishra, Sanja Petrovic, and Santhanam Sundar. "A self-adaptive case-based reasoning system for dose planning in prostate cancer radiotherapy". *Medical physics* 38.12 (2011), pp. 6528–6538.
- [67] Judea Pearl. *Probabilistic inference in intelligent systems*. 1988.
- [68] Peter JF Lucas, Linda C van der Gaag, and Ameen Abu-Hanna. "Bayesian networks in biomedicine and health-care". *Artificial intelligence in medicine* 30.3 (2004), pp. 201–214.

-
- [69] Charles E Kahn, Linda M Roberts, Katherine A Shaffer, and Peter Haddawy. "Construction of a Bayesian network for mammographic diagnosis of breast cancer". *Computers in biology and medicine* 27.1 (1997), pp. 19–29.
- [70] PJF Lucas, CAM Schurink, IM Hoepelman, and MJM Bonten. "A Model-based Approach to Improved Prescription of Antibiotics" (2003).
- [71] Athanasios Tsoukalas, Timothy Albertson, and Ilias Tagkopoulos. "From data to optimal decision making: a data-driven, probabilistic machine learning approach to decision support for patients with sepsis". *JMIR medical informatics* 3.1 (2015), e11.
- [72] HJ Gómez-Vallejo, B Uriel-Latorre, M Sande-Meijide, B Villamarín-Bello, Reyes Pavón, F Fdez-Riverola, et al. "A case-based reasoning system for aiding detection and classification of nosocomial infections". *Decision Support Systems* 84 (2016), pp. 104–116.
- [73] Bente Arboe, Rasmus Rude Laub, Gitte Kronborg, and Jenny Dahl Knudsen. "Evaluation of the decision support system for antimicrobial treatment, TREAT, in an acute medical ward of a university hospital". *International Journal of Infectious Diseases* 29 (2014), pp. 156–161.
- [74] Timothy M Rawson, Luke SP Moore, Bernard Hernandez, Esmita Charani, Enrique Castro-Sanchez, Pau Herrero, et al. "A systematic review of clinical decision support systems for antimicrobial management: are we failing to investigate these interventions appropriately?" *Clinical Microbiology and Infection* (2017).
- [75] Karin Thursky. "Use of computerized decision support systems to improve antibiotic prescribing". *Expert review of anti-infective therapy* 4.3 (2006), pp. 491–507.
- [76] Christopher E Curtis, Fares Al Bahar, and John F Marriott. "The effectiveness of computerised decision support on antibiotic use in hospitals: A systematic review". *PloS one* 12.8 (2017), e0183062.
- [77] Elisa G Liberati, Francesca Ruggiero, Laura Galuppo, Mara Gorli, Marien González-Lorenzo, Marco Maraldi, et al. "What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation". *Implementation Science* 12.1 (2017), p. 113.
- [78] Annette Moxey, Jane Robertson, David Newby, Isla Hains, Margaret Williamson, and Sallie-Anne Pearson. "Computerized clinical decision support for prescribing: provision does not guarantee uptake". *Journal of the American Medical Informatics Association* 17.1 (2010), pp. 25–33.
- [79] Zhiping Walter and Melissa Succi Lopez. "Physician acceptance of information technologies: Role of perceived threat to professional autonomy". *Decision Support Systems* 46.1 (2008), pp. 206–215.
- [80] Cynthia S Gadd, Prakash Baskaran, and David F Lobach. "Identification of design features to enhance utilization and acceptance of systems for Internet-based decision support at the point of care." *Proceedings of the AMIA Symposium*. American Medical Informatics Association. 1998, p. 91.
- [81] Viswanath Venkatesh, Michael G Morris, Gordon B Davis, and Fred D Davis. "User acceptance of information technology: Toward a unified view". *MIS quarterly* (2003), pp. 425–478.

- [82] Saif Khairat, David Marc, William Crosby, and Ali Al Sanousi. "Reasons For Physicians Not Adopting Clinical Decision Support Systems: Critical Analysis". *JMIR medical informatics* 6.2 (2018).
- [83] Martin Hoenigl, Reinhard B Raggam, Jasmin Wagner, Thomas Valentin, Eva Leitner, Katharina Seeber, et al. "Diagnostic accuracy of soluble urokinase plasminogen activator receptor (suPAR) for prediction of bacteremia in patients with systemic inflammatory response syndrome". *Clinical biochemistry* 46.3 (2013), pp. 225–229.
- [84] Tsipi Heart, Ofir Ben-Assuli, and Itamar Shabtai. "A review of PHR, EMR and EHR integration: A more personalized healthcare and public health policy". *Health Policy and Technology* 6.1 (2017), pp. 20–25.
- [85] Nitika Pant Pai, Caroline Vadnais, Claudia Denking, Nora Engel, and Madhukar Pai. "Point-of-care testing for infectious diseases: diversity, complexity, and barriers in low-and middle-income countries". *PLoS medicine* 9.9 (2012), e1001306.
- [86] Manuela Ölmez and Udo Lindemann. "Managing attribute complexity for user-centered decision support systems". *Procedia Computer Science* 28 (2014), pp. 130–137.
- [87] Pieter J Beers and Wim Gijsselaers. *Decision-support and Complexity in Decision Making 1 Decision-support and Complexity in Decision Making ACADEMIC TRACK*. Tech. rep.
- [88] Timothy Miles Rawson, Esmira Charani, Luke Stephen Prockter Moore, Bernard Hernandez, Enrique Castro-Sánchez, Pau Herrero, et al. "Mapping the decision pathways of acute infection management in secondary care among UK medical physicians: a qualitative study". *BMC medicine* 14.1 (2016), p. 208.
- [89] Robert James Campbell. "The five rights of clinical decision support: CDS tools helpful for meeting meaningful use". *Journal of AHIMA* 84.10 (2013), pp. 42–47.
- [90] Lise Poissant, Jennifer Pereira, Robyn Tamblyn, and Yuko Kawasumi. "The impact of electronic health records on time efficiency of physicians and nurses: a systematic review". *Journal of the American Medical Informatics Association* 12.5 (2005), pp. 505–516.
- [91] Rochelle Brooks and Courtney Grotz. "Implementation of electronic medical records: How healthcare providers are managing the challenges of going digital". *Journal of Business & Economics Research* 8.6 (2010), pp. 73–84.

CDSS DESIGN AND IMPLEMENTATION

This chapter describes the clinical decision support system (CDSS) implemented during the conducted research. In the first place, the overall system architecture is explained (section 3.1), paying particular attention to the integration with various electronic health databases, the accessibility at point of care and the modular design methodology. To continue, the chapter explains the modules providing decision support (section 3.2), the design strategies (subsection 3.3.1) and the iterative process used to evaluate the usability of the system (subsection 3.3.2). After this, an illustration of the graphical user interface for each of the modules is presented (section 3.3). To conclude, the chapter discusses the main areas in which the CDSS contributes to provide effective care.

3.1 DECISION SUPPORT SYSTEM OVERVIEW

Enhanced, Personalized and Integrated Care for Infection Management at Point of Care (EPiC IMPOC) is a decision support system designed to record a complete set of vital signs at the patient's bedside on hand-held computing devices while providing instant decision-making assistance to health care workers at the point of care. It is integrated within the electronic health records and therefore has access to the hospital patient administration system and laboratory data. It is accessible anywhere in the hospital by members with appropriate access rights through mobile devices or desktop computers connected to the hospital intra-net. The CDSS architecture (see Figure 3.1) can be divided into three main components: the external sources of patient data, the server application with the decision support modules (see section 3.2) and the client application with the graphical user interface (see section 3.3).

■ *Integration with electronic health records*

The integration of any CDSSs into health care information technology systems facilitates the collection of data to provide effective and efficient health care. As such, EPiC IMPOC feeds from three databases within the National Health Services (NHS). Firstly, the patient administration system provides demographic data such as age or gender and hospital management information such as insurance number, date of admission or admitted ward. The pathology system contains biochemical laboratory results such as bilirubin, c-reactive protein and white cell counts. Lastly, the microbiology system provides susceptibility test outcomes to identify the presence of pathogens and determine which antimicrobials are effective. These include specimen collection source, infectious pathogen, antimicrobial and sensitivity outcome. These databases will be thoroughly explained in following chapters.

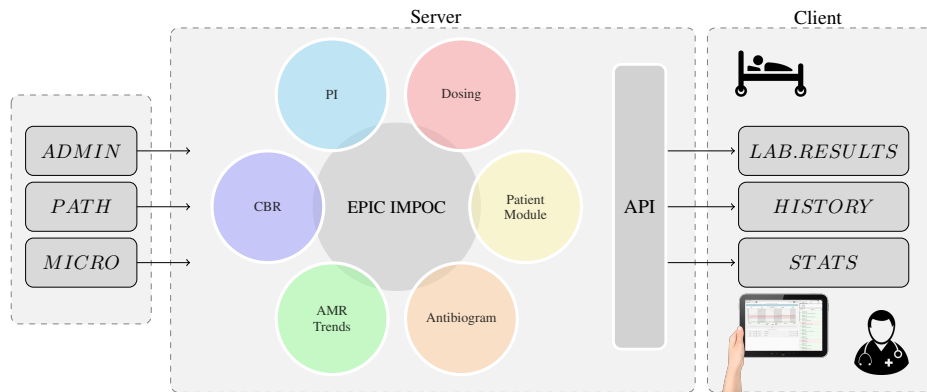


Figure 3.1: EPiC IMPOC: high-level system architecture diagram. It describes the main components of the CDSS. The external databases are the patient administration system (ADMIN), pathology laboratory tests (PATH) and microbiology results (MICRO). The server application has the following modules: case-based reasoning (CBR), probabilistic inference (PI), patient engagement module, therapeutic drug monitoring and AMR surveillance. The information is accessed through an API and it is presented on hand-held computer devices or desktop computers to clinicians.

- *Accessibility at point of care*

Access to point-of-care (POC) tools has the potential to enhance rapid disease diagnosis, management, control and surveillance [1]. Furthermore, providing decision support information at POC facilitates communication among health care workers (e.g. nurses, clinicians and infection specialists) and shortens the time for clinical decision making to either plan further investigations (e.g. laboratory or susceptibility testing) or select the antimicrobial therapy. These POC services are especially useful in settings where access to quality and timely medical care is a challenge.

- *Modularity in design*

A modular system can be characterized by functional partitioning into discrete scalable and reusable modules. This design approach is used primarily to reduce complexity by breaking a system to encapsulate the desired functionality of each component behind a module and its interface [2]. As shown in Figure 3.1, EPiC IMPOC has the following decision support modules: patient engagement, antimicrobial dosing, antimicrobial surveillance (AMR trends and antibiogram), case-based reasoning and probabilistic inference.

- *Software tools*

The decision support system was implemented using Java in the server side and HTML, CSS and Javascript for the user interface. The jColibri library [3] was used to implement case-based reasoning within the server. The Python programming language has been used for machine learning prototyping. In particular, supervised learning models and performance metrics from Scikit-learn [4] and sampling techniques from Imbalanced-learn [5] were employed. Data handling was done with Pandas [6, 7] and data visualization using Matplotlib [8] and Seaborn [9].

3.2 PART I: OUTLINE OF DECISION SUPPORT MODULES

The server application processes queries, interacts with the permanent storage and provides decision support. It has been implemented in Java and uses an object relational mapping java library (Hibernate ORM) to map an object-oriented domain model to a traditional relational database (SQL). The Lightweight Directory Access Protocol (LDAP) accomplishes the authorization and authentication of users and it is provided in all hospitals at Imperial College Healthcare National Health Service Trust. The server implementation follows the Representational State Transfer (REST) architectural design [10], which suggests a group of guidelines to create scalable, more performing and maintainable web services. A complete description of the application programming interface (API) including subroutine definitions and communication protocols is provided as supplementary material (see Appendix C). The modules composing EPiC IMPOC are briefly explained below.

- *Promoting patient engagement with antimicrobial decision making*
This module generates a single document summarising the key aspects of the infection (name, site and brief description) and the selected antimicrobial therapy (antimicrobials, dosage and course of treatment). In addition, it provides official sites to access further related information. The aim of the module is to promote patient engagement with decision-making to improve their understanding of antimicrobial use and shape future patient behaviour [11, 12].
- *Antimicrobial dosing and therapeutic drug monitoring*
This module implements therapeutic drug monitoring guidelines to enhance drug efficacy while reducing toxicity on the patient. As such, periodic measurements of drug concentrations need to be taken to individualise dosage to maintain concentrations within a target range or reach a particular peak [13]. Drug concentration at the site of infection cannot be routinely measured, but the desired or adverse effects may correlate better with plasma or blood concentration than they do with dose.
- *Antimicrobial surveillance to identify local resistance patterns*
In the last years, a variety of initiatives have arisen to promote prudent use of antimicrobials and enhance global AMR surveillance [14, 15]. Despite of these initiatives, homogeneity of antimicrobial policies still produces different AMR outcomes [16]. For example, resistance rates are higher in London than in the rest of UK due to complex reasons [17–20]. This module computes hospital-specific AMR surveillance to promote education and awareness of local resistance patterns (see Chapter 4). Furthermore, this information is also valuable to support antimicrobial therapy selection.
- *Case-based reasoning for inspection of previous similar cases*
This module facilitates inspection of past similar cases to promote evidence-based medicine, support antimicrobial therapy selection and enhance interpersonal communication (see Chapter 5). For such purpose, the case-based reasoning methodology has been implemented to retrieve from the database previous cases describing patients with similar symptoms, the antimicrobial therapy applied and the corresponding outcome. It is worth noting that this knowledge-based methodology, which closely replicates the cognitive process followed by health care workers, has strongly influenced the design of the user interface.

■ *Probabilistic Inference to provide stepwise decision support*

The acute infection management pathway followed by clinicians has been defined as a stepwise Bayesian model in which each step adds systematically information to optimise diagnosis and management of infection [21]. This module focuses on two steps: (i) estimate the likelihood of infection using culture positivity as proxy and (ii) anticipate the most plausible sites of infection. This module aims at reducing unnecessary antimicrobial prescriptions, refine empirical therapies and guide suitable specimen collection for further susceptibility testing (see Chapter 6).

3.3 PART II: THE GRAPHICAL USER INTERFACE

Medical knowledge is voluminous and changes rapidly. For instance, vital signs of a patient might change unexpectedly and physicians need to be aware of these variations to make correct medical decisions. CDSSs are often devised as a potential solution to aid clinicians through the organization and analysis of such voluminous information. However, CDSS adoption remains low [22]. To promote the use of CDSSs by physicians, it is fundamental to determine the best way to present the information via interfaces [23]. User interfaces can be designed in several ways [24]. In one strategy, designers try to match the design of the interface with the decision process so as to facilitate physician navigation [25]. Alternatively, designers may focus on usability principles [26, 27]. In this research, both strategies have been combined to design EPiC IMPOC.

3.3.1 User interface design strategies

The design of user interfaces according to the decision process is generally read from top to bottom. In the first section, these interfaces include data entries and check boxes for physicians to collect the necessary data [28, 29]. Note that data collection might take time [30]. After this, physicians press buttons to go through the decision process. To navigate through the different sections, the use of expand/contract interfaces is also common [31, 32]. Many of these expand/contract systems depend on the length of the hierarchy, with no complete view of the overall arborescence. In EPiC IMPOC, the expand/contract sections have been fixed so as the number of clicks is always the same and clinicians have a priori knowledge on the navigation system. To conclude, highlight the importance of minimizing the number of screens to improve navigation [33–35].

The design of user interfaces according to usability principles relies on a set of well defined techniques. First, it is imperative to use concise, consistent and unambiguous language to promote simple and efficient interactions [33, 34]. Also, explanations and justifications might be presented to increase physician confidence [33]. These interfaces should provide advice and suggestions but also alternatives to increase compliance and respect the autonomy of the physician [33]. The number of clinical parameters used by physicians is large and diverse. Hence, the most relevant information needs to be placed in more prominent positions to ensure that it is seen [33, 34]. It is also a good practice to group similar pieces of information together to facilitate on-screen searches [24]. Elements such as tables, graphs, buttons, scroll bars or iconic language might help reducing the perceived density of the interface [33, 35].

To conclude, highlight basic display aspects within interface design such as appropriate font sizes, adequate contrast between text and background and meaningful colours to enhance readability [34, 35].

3.3.2 Overview of the design process

In this research, the design of EPiC IMPOC has considered the previously described interfaces which are based on both usability principles and the physician decision process. A summary of the iterative methodology used to develop and evaluate the interface is presented in Table 3.1. In the first iteration, the interface was designed to align with the case-based reasoning methodology (see Chapter 5). As such, the main focus was to facilitate homogeneous collection of vital signs, enhance the review and monitoring of the current patient and augment further investigation of past similar cases. A paper-based prototype was presented to two infection specialists (see LM and TR in Table 3.1) who were collaborating in this research. The feedback was positive and the first web-based prototype was implemented using HTML, CSS and Javascript. As such, it was accessible through any device using a web browser. The implementation followed a responsive design approach to render a consistent interface across different devices and screen sizes from desktop computers to hand-held devices.

Table 3.1: EPiC IMPOC summary of the design process.

	Description	Feedback	Outcome
Iteration 1	The paper-based prototype was described to LM and TR (infection specialists).	- Overall positive feedback	
Iteration 2	The web-based prototype was trialled in the ICU for a month led by LM and TR to evaluate the CBR module (see subsection 5.5.1).	- Improve terminology and consistency - Inclusion of new relevant vital signs - Organization of vital signs within the form - Identification of decision process (5 steps) - 100% data retrieval fidelity	[36]
Iteration 3	A survey to evaluate the usability of the web-based prototype using the SUS form was conducted (see subsection 5.5.2). The evaluation considered exclusively the CBR module.	- Reduce the amount of information - Provide specific decision support - Fit within the infection management pathway	[21, 37]

Keys: CBR=case-based reasoning; ICU=intensive care unit; PI=probabilistic inference; SUS=system usability scale; LM=Luke Stephen Prockter Moore, TR=Timothy Miles Rawson;

Note: LM and TR are infection specialists who collaborated in this research.

The web-based prototype was trialled in the ICU for a month (see subsection 5.5.1). Feedback was provided as to the terminology that should be used, the organization of the vital signs within the data collection form and the overall decision process. As a result, the decision process was divided into five steps: data collection, pathology review, microbiology investigations, inspection of previous similar cases and antimicrobial therapy selection. Moreover, during this period the fidelity of the data collected by the system was 100%. The feedback provided was included into the prototype and a survey to evaluate the usability of the CDSS interface was conducted (see subsection 5.5.2). The survey was completed by 10 health care workers including nurses, physicians and infection specialists and the overall feedback was positive [37]. They highlighted three

main areas for improvement: (i) the vast amount of data presented is useful for experts but might lead to confusion if used by inexperienced physicians (ii) the need of more specific decision support to better guide inexperienced physicians and promote its use under time constraints and (iii) the need to fit within the infection management pathway followed by nurses and physicians to prescribe antimicrobials.

Further research focused on understanding the infection management pathway followed by clinicians [21]. The reported pathway was defined as a stepwise Bayesian like approach in which each step adds systematically information to optimise diagnosis and management of infection. Thus, the probabilistic inference module (see Chapter 6) was included in EPiC IMPOC to provide support on each of the previously identified steps. The overall feedback was very positive with all participants considering that the implementation and use of the system in clinical practice is a good idea.

3.3.3 Illustrating the medical record of a patient

The main unit of information in the system is the medical record of a patient (see CBR case in Table 3.2). In the first instance, a case only shows a brief synopsis of the patient condition and the antimicrobial therapy prescribed (C1). Once a case is selected, the complete medical record is shown in five different tabs: resume, description, pathology, sensitivity and solution. The resume section (C2) displays the symptoms and the description section (C3) enables further modification and completion of parameters within the patient case. To evaluate the progression of the patient, the temporal evolution of some biochemical markers such as bilirubin or white blood cell counts are shown in the pathology section (C4). They are represented as graphical time series where coloured background indicates the normal reference range. Additionally, it is possible to hide/show time series to improve visualization. The microbiology section (C5) contains the susceptibility tests outcomes available for specific pathogen-antibiotic combinations grouped in three categories: resistant, intermediate and sensitive. These outcomes are also presented during therapy selection for further guidance. Finally, the solution section (C6) contains the antimicrobial therapy prescribed (drug, dose and route of administration) and further feedback from clinicians.

Table 3.2: EPiC IMPOC: the medical record of a patient representing a CBR case.

C1 2690 2MAY15 ♀ Vancomycin C2 Status Today Pathology Microbiology Outcome INFECTION-INFO Organs Inf. Urine Bugs ECOL INFECTION-MARKERS Isotopes null Body Temp. 37.5 CL Changed false ABDOMINAL Abdominal Exam. Tender Renal Supp. HDX Urinary Cath. true RESPIRATORY-SYSTEM Resp. Rate 17 Oxygen Req. 21 Ventilation Supp. Own Chest Rad. Clear Chest Exam. Clear DEMOGRAPHICS HR 65 Blood Pre. 102 Height 155 Weight 55 Pregnant ☒ Diabetes ☐ HIV ☐ Allergies Aminoglycosides Antifungals Echinocandins Glycopeptides Macrolides Save
--

An example showing the graphical representation of a patient case within the GUI. The patient case is divided in six sections: synopsis, resume, description, pathology, microbiology and solution. On the first instance, exclusively the synopsis is visible. By clicking on the synopsis the patient clinical record with five tabs is displayed. These tabs are browsed by either clicking the tab name or swiping (left/right).

3.3.4 Understanding the main control panel

The main control panel of the graphical user interface (GUI) has four areas clearly differentiated (see Table 3.3): the navigation menu, patient selection, main dashboard and sidebar. Firstly, a patient can be selected (D1) by introducing either NHS or hospital number in the search box or by choosing a patient from the drop down list. Once the patient is selected, the most recent case (D2) and previous medical records (D3) are automatically retrieved. In addition, the sidebar which contains three tabs (*inference*, *cases* and *groups*) is automatically populated. The aim is to reduce the number of interactions between the clinician, the interface and the server (see basic interaction diagram in Figure 3.2).

The side bar is composed of three tabs: *inference*, *cases* and *groups*. By default, the *inference* tab which provides stepwise decision support (D6) is displayed. The aim of this tab is to provide systematic information of the infection status of the patient under investigation. On the top, it presents the most recent laboratory test results available for the selected biochemical markers. Below, the risk of infection and the most plausible sites are presented. The other two tabs (*cases* and *groups*) display previews of past similar cases which have been retrieved from the database. The *cases* tab (D5) contains a list of case previews which can be sorted/grouped by any of the parameters composing the clinical record of a patient. The *groups* tab classifies such case previews by outcome (success/failure) and prescribed antimicrobial therapy. In addition, it displays the number of case previews in which the therapy

was applied and the average similarity score. By clicking on these previews, the corresponding cases are added to the main dashboard for further investigation (D4). The vertical bar on the left hand side of the case preview indicates that the clinical record associated is currently being displayed in D4. The side bar can be hidden to facilitate inspection of clinical cases by clicking on the icon on the top right corner. To conclude, all the cases displayed within D2, D3 and D4 include all the information (resume, pathology time series, microbiology results, antimicrobial therapy prescribed and outcome) of a patient clinical record as explained previously in section 3.3.3. Note that those cases with the lock symbol in the synopsis cannot be modified.

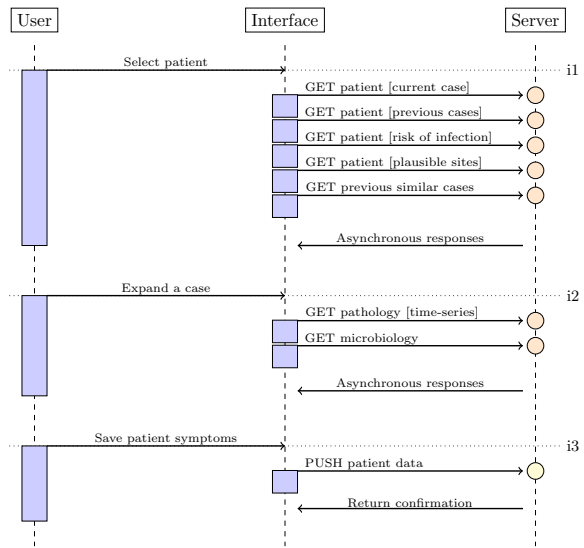


Figure 3.2: EPiC IMPOC main panel: basic interaction diagram. Three basic interactions are demonstrated in the diagram: (i) the selection of a patient which retrieves automatically the current case, the previous cases, past similar cases, the probability of infection and the most plausible sites (ii) the expansion of a case which automatically pulls the pathology (time series) and microbiology data and (iii) the collection of symptoms which are stored in the database. Orange circles indicate read access to the database. Yellow circles indicate writing access to the database.

D1

D2

D3

D4

D6

D5

The clinical record of an inpatient can be selected by introducing NHS or hospital number in the search box or choosing a patient from the drop down list which can be previously filtered by ward.

This section shows the most recent clinical record for the selected patient so vital signs can be reviewed and/or updated appropriately.

This section shows previous clinical records for the selected patient which can be the current or previous admissions within the hospital.

This section contains the clinical records of previously admitted patients with similar symptoms for further clinical inspection. The cases displayed in this section are added/removed by clicking on the case preview within the side bar (D5). Each of the cases presented in this section has all the information of a patient clinical record as explained in section 3.3.3.

This section shows a preview of the retrieved clinical records with the similarity score, the successful score, the applied antimicrobial therapy and the outcome (background colour). By default, the case previews are ordered by similarity in descending order. However, they can be sorted/grouped by any of the attributes. The vertical bar indicates that the clinical record is being displayed in D4 and can be added/removed by clicking on the case previews within the side bar (D5).

Firstly, this section presents the most recent laboratory test results available for the selected biochemical markers. These markers are also used to estimate the risk of infection and the most plausible sites

In future work, this tab would also provide a summary of scores/probabilities describing relevant patient conditions such as systemic inflammatory response syndrome (SIRS) or chronic obstructive pulmonary disease (COPD). Following the step wise decision support, it will include most likely pathogens (species and/or genera) causing the infection and antimicrobial sensitivity maps for such pathogens (see Chapter 4).

An example showing the main control panel of the GUI. The sidebar contains three tabs: inference, cases and groups which are browsed by clicking on the corresponding tab name. The top left icon allows to select other views such as a tutorial or the AMR surveillance.

3.3.5 Design techniques and usability principles within EPiC IMPOC

The link between the design techniques [33, 35] included in EPiC IMPOC and the usability principles [26, 35] are summarised in Table 3.4. Firstly, the main control panel has been organised into fixed areas (see Table 3.3) to facilitate on-screen searches. These areas are patient selection (D1), patient clinical record (D2 and D3), past similar cases (D4) and the side bar (D5). In addition, the most important information such as the patient clinical record (D2) or the stepwise decision support (D5) is prominent to ensure that it is seen. For instance, the patient clinical record is on the top left and the stepwise decision support section is displayed by default. In order to align with the usability principles, similar pieces of information are grouped together. Thus, the information is organised in patient cases (see Table 3.2) whose structure is consistent to improve readability. Moreover, each case is in sections so as to display parameters extracted from the same source together. These sections are the synopsis of the patient case (C1), resume (C2), description (C3), pathology (C4), microbiology (C5) and solution (C6). The synopsis of the patient case (C1) uses iconic language to display relevant information succinctly. In the description section, data entries and check boxes for physicians to collect the vital signs have been included. This is common in interfaces designed according to the decision process. Note that data collection might take time [30] and was a common concern among the participants. The aim of EPiC IMPOC is to collect all the vital signs automatically from the NHS servers, yet access was not provided during this research. In addition, the pathology data (C4) and microbiology data (C5) are presented in a graph and a table respectively to facilitate visualization.

The side bar (D5) contains three sections: one with the step wise decision support (which is shown by default) and two with different previews of the CBR cases. These sections are read from top to bottom to minimize cognitive load and promote naturalness. For instance, the CBR cases retrieved from the database are sorted by similarity. Moreover, meaningful colours have been used consistently. In the step wise decision support section, a range of colours from yellow to red are used to represent the severity of the infection. In the CBR case preview section, the therapy outcomes are presented in green when they were a success or red when they were not effective. The use of meaningful colours promotes effective information presentation and improves readability. Moreover, EPiC IMPOC provides advice and suggestions, but also displays the alternatives to increase compliance and respect the autonomy of the physician [33].

Note that the use of expand/contract interfaces maximizes the amount of information presented. Many of these expand/contract systems depend on the length of the hierarchy, with no complete view of the overall arborescence. In EPiC IMPOC, the expand/contract sections have been fixed so as the number of clicks is always the same and clinicians have a priori knowledge on the navigation system. For instance, all cases can be either expanded/contracted and the information within the case is consistent. To conclude, the visualization is presented on a single screen so as to preserve the context. This space-filling approach leads to an increase in the amount of information presented in the available space. While participants reported that the interface is a bit busy, experienced infection specialists who were familiar with the interface highlighted the advantage of having all the information available during antimicrobial therapy selection. Hence, providing a training session to future users would potentially reduce the perceived complexity of the system while ensuring an appropriate use.

Table 3.4: Link between the design techniques [33, 35] and usability principles [26, 35] used to build the EPiC IMPOC interface.

Design technique	Effect on visualization	Simplicity	Naturalness	Consistency	Minimize cognitive load	Efficient interactions	Forgiveness and feedback	Effective use of language	Effective information presentation	Preservation of context
Space-filling approach.	The amount of information displayed in the available space is maximized.								X	X
In addition to suggestions, alternatives are presented during the decision process.	Provides an overview of the decision process and preserves the autonomy of the physicians.						X		X	
Organization into fixed areas. Each area contains always the same kind of information.	On-screen searches are facilitated because physicians know where to find the information they need.	X		X	X	X	X		X	
Localization of information. The most relevant data is prominent and visible.	The most important information is sure to be read by physicians.				X	X			X	
Group similar pieces of information together.	On-screen searches are facilitated and cognitive load reduced.		X	X	X	X			X	
Decision process read from top to bottom.	Readability is improved.		X		X	X			X	
Use of a consistent representation of a case.	Readability is improved.	X		X	X					
Use of iconic language.	Readability is improved.	X	X						X	
Use of graphs and tables.	Readability is improved.	X	X						X	
Use of meaningful colours.	Readability is improved.	X	X	X		X			X	
Use of concise language to describe decision and action variables.	Readability is improved.	X						X		
Use of acceptable contrast between text and background.	Readability is improved.		X							
Visualization in a single screen	Readability is improved								X	
Expand/contract interfaces.	Readability is improved					X			X	X

3.3.6 *The patient engagement module*

In order to understand patient engagement with decision making in infection management a qualitative investigation of current experiences was undertaken [11]. The investigation was led by Timothy Rawson. A total of 10 members of the public who had received antimicrobials from secondary care in the preceding 12 months in the UK participated in the study. The group was divided into two equal groups based on age categories and gender. The median age of participants was 53 (21-69) with equal gender proportions. Focus groups were audio recorded and three main themes were identified: communication, information provision and media. The participants reported a lack of communication and information provision from clinicians which materialised in feelings of dis-empowerment among participants and led to loss of ownership, frustration and in some cases anxiety. This potentially alters the future actions of patients towards infections and antimicrobials. Moreover, it can drive non-adherence to prescribed antimicrobial regimes and loss to follow-up after discharge from secondary care.

In response, a personalised antimicrobial information module co-designed with patients was developed [12]. The study evaluated the potential impact on short-term knowledge and understanding of antimicrobial therapy in secondary care. A total of 30 previous patients who had received antibiotics in secondary care during the preceding 12 months participated in two separate one hour workshops. The first workshop was held in September 2015 and the second in May 2016. There was consistency in the themes that emerged in both workshops. First, the platform needed to be flexible and personalised. Participants agreed upon the development of a personalised PDF document containing data specific to the individual. Moreover, they reported that the PDF was the optimal approach as it allowed the maximum flexibility to either be printed and given to a patient at the bedside or transferred electronically. Other approaches considered included the development of a mobile application, text message services and written summaries. Second, the intervention should provide information about the individual's condition and treatment where the information must be provided in language that the majority of the citizens can understand. They reported that detailed descriptions should not be included, but references for reputable sources of information should be provided. Finally, the intervention should provide practical advice on the side effects and serve as tool to enhance communication and support follow up.

As a result, a patient engagement module that provides personalised antimicrobial information in the form of a leaflet was implemented in EPiC IMPOC. The final template that was agreed upon and co-design by participants in the workshops (see sections A1-A6 in Table 3.5) is automatically populated and presented in a PDF document which can be either printed or mailed. A number of translations occur during the leaflet generation. For instance, if the technical term pneumonia is recorded by the physician, the term is coded as chest infection. Moreover, this translation also triggers the inclusion of a web address with further open access information on pneumonia. Therefore, on generating the information leaflet through the clinical decision support tool, the clinician is able to provide a personalised information leaflet to the patient, which contains details of their own infection and treatment. Although embedded in an electronic clinical decision support system that contains several different modules, the intervention currently sits independently of these.

Table 3.5: EPiC IMPOC: the patient engagement report.

A1	Doctor: Unknown Unknown (2) Beep: Not available 14 February 2017	Doctor information [A1]	Patient information [A2]
A2	Hospital Number: 30256719 Name: Padme Amidala Date of Birth: 1945-07-01	This section presents the name of the doctor, a contact number and the date in which the patient engagement report has been produced.	This section presents the the following patient information: hospital number, full name and date of birth.
A3	Your doctor thinks that you have the following infections: This can also be called: Abdomen - http://patient.info/doctor/intra-abdominal-sepsis-and-abscesses		
A4	Your doctor thinks that a bacteria may be causing your infection. - Bacteria found in your laboratory tests: Acinetobacter Baumannii	Clinical suspicion [A3]	Infectious pathogen [A4]
A5	- Treatment received previously (starting on None): - They have now prescribed you the following antibiotic treatment (on 02 May 2015): Ciprofloxacin ... by mouth (orally) every 12 hours ... for 5 days. Metronidazole ... by mouth (orally) every 8 hours ... for 5 days. Sulbactam ... through the vein (intravenously) every 24 hours ... for 3 days.	This section presents the suspicion of infection provided by the physician and external links for further information.	This section displays the infectious pathogen. This section it is filled in those scenarios in which the microbiology requested were positive.
	- Commonly reported side effects Allergy Nausea Discomfort - You CANT drink alcohol or drive whilst taking these antibiotics Always read the information leaflet before taking any antibiotics.		
A6	Access my results: You can sign up to view your own medical information and access your test results through the following free website, called "Patient Knows Best". http://www.patientknowsbest.com Drug resistant bacteria can grow if you do not take antibiotics as they are prescribed. This may mean that you do not get better from your infection and can make future infections more difficult to treat. To prevent the development of drug resistant bacteria always take your course of antibiotics as prescribed by your doctor. If you have concerns or are unable to complete the course you should discuss this with a healthcare professional. Further information on drug resistant infections and support with taking antibiotics can be found at: http://www1.imperial.ac.uk/epicimprovement/patient_and_public/	Antimicrobial therapy [A5]	Further links [A6]
		This section presents the past and current antimicrobial therapies including the prescribed antimicrobials, the route of administration, the frequency and the course of treatment. In addition, the leaflet presents a series of commonly reported side effects.	This section presents a variety of links to access information such as their pathology laboratory tests results. Moreover, it provides web addresses that directly link to open access information related with infectious diseases, antimicrobial resistance and support with taking antibiotics.

An example showing the patient engagement report generated where data is automatically retrieved from the database. An investigation of current patient engagement in secondary care and a further intervention were conducted to identify the relevant parameters/format and evaluate the effects respectively [11, 12].

A posterior intervention with pre- and post-intervention questionnaires was piloted over a month on 30 patients in which fifteen out of thirty (50%) completed the study [12]. Pre-intervention, participants reported the need of more information about their infections and the antibiotic therapy than they had been given. Post-intervention, participants reported that the information provided was useful and had not yet been given to them by their doctor (e.g. information about their infection, the antimicrobial therapy or general side effects). Feedback provided by patients highlighted that this intervention prompted more detailed discussions with the doctor. Overall, the study highlighted the poor baseline knowledge of antibiotic therapy and infection management amongst hospitalised patients being treated for infections and how the provision to patients of a simple personalised leaflet supports and improves short term knowledge and understanding in secondary care.

3.3.7 The AMR surveillance module

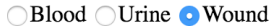
The graphical user interface developed for this module (see Table 3.6) was created with the sole purpose of visualizing and evaluating the strategies presented to enhance AMR surveillance (see Chapter 4). These strategies compute AMR-related statistics such as antimicrobial resistance rates (see section 4.1), antimicrobial spectrum of activity (see section 4.5) or antimicrobial resistance trends (see section 4.6). For further integration of these results into EPiC IMPOC, a qualitative investigation to select the main components, the preferred visualization methods and the appropriate strategy to integrate such information into the CDSS so as to be presented according to the infection management pathway needs to be conducted.

To begin with, a tree map with pathogen-antimicrobial pairs is presented (B1). Pathogens are represented by colours and antimicrobials by their acronyms. Thus, each rectangle corresponds to an specific pathogen-antimicrobial pair. The size of the rectangle represents the proportion of susceptibility tests for a specific pair. For instance the combination *Escherichia coli* (colour blue) and erythromycin (ERY) is the most tested. By hovering on the rectangle the total number of susceptibility tests available is displayed. The susceptibility tests have been grouped by culture site and the three most common (blood, urine and wound) have been presented. A number of radio buttons are provided below to visualize the tree map for a particular site. Once a rectangle is clicked, the corresponding image is displayed on the right with information related to the selected pathogen-antimicrobial pair. First, on the top left (B3) the resistance time series and the number of susceptibility tests available are displayed. Below, various statistical tests (B2) to describe the time series are presented. These tests determine the approximated trend (theilsens slope), the monotony of the time series (kendall test) and the stationariness of the time series (adfuller test). This information is useful to discern which methods for time series analysis can be applied (see Chapter 4). In addition, this area (B2 top) shows the correlation between the number of susceptibility tests available and the computed resistance index. Note that the resistance index should not be correlated with the number of susceptibility tests. Correlation between the two might indicate that the number of susceptibility tests available is insufficient.

In order to estimate the antimicrobial resistance trend a number of regression models have been considered (see section 4.6). A summary of the results for ordinary least squares (B4 bottom), weighted least squares (B4 top) and some autoregressive integrated moving average models (B6) are also presented. These summaries include the name of the method, the coefficients of the model and a number of metrics to evaluate the performance (Durbin-Watson or Jarque-Bera). The graphical representation of these models is provided in B5. Further information such as temporal autocorrelation or time series decomposition into trend, seasonal component and noise are also displayed B7.

All the information have been saved in a csv file and it is being displayed as a single pdf figure at the moment. Firstly, a qualitative investigation should be conducted to identify the AMR statistics that are relevant for physicians and therefore need to be included into EPiC IMPOC. Moreover, since EPiC IMPOC is a web-based application, further integration using a dynamic, interactive and data-driven visualization tool is recommended¹.

¹<https://d3js.org>



[B1]

[B2]

[B3]

[B4-B7]

An example showing the partially interactive antimicrobial resistance module. The presented interface was used to research trend estimation and forecasting. Therefore, further work should focus on identifying the relevant parameters for doctors to display them through a data-driven interactive visualization tool.

3.4 DISCUSSION

An appropriately designed clinical information technology system has the potential to assist clinicians in many different aspects such as antimicrobial therapy selection [38–40] or patient monitoring [41]. The majority of CDSSs focus on a single aspect of the infection management pathway hence limiting their adoption in clinical practice. In contrast, EPiC IMPOC assists in various aspects of the infection management pathway to maximize the system utility.

Improvement of consistency in vital sign data collection

The efficiency of any CDSSs depends on the quality of the underlying data. Existing strategies for vital sign data collection rely on paper-based forms resulting in low currency, completeness and interoperability [42]. Furthermore, such forms are often manually transcribed into electronic records. To simplify the data collection process and improve homogeneous collection of vital signs, a standardised digital documentation form (see C3 in Table 3.2) is provided at the point of care.

Facilitate patient history review and monitoring

Health care workers have to integrate several physiologic parameters to evaluate the state and detect changes in patients. The high cognitive demand for data integration reduces the available resources for communication, documentation or decision making to take corrective actions which can lead to a cascade of errors. Therefore, providing information about the status of the patient in a manner that is easy and fast to interpret reduces the time needed to observe changes and therefore promotes early detection of adverse events.

Facilitate review of previous similar cases

The cognitive process followed by clinicians to prescribe antimicrobials has three main components: evaluate the symptoms of the patient in front of them, identify similar situations based on their own clinical experience and use that information to select the appropriate antimicrobial therapy [21, 43]. In many cases, such as junior doctors, this experience is somehow limited [37]. Therefore, the integration of case-based reasoning to retrieve previous similar cases (see D4 and D5 in Table 3.3) empowers clinicians through the creation of a shared knowledge base from which cases are retrieved in a coherent fashion and presented seamlessly at point of care to augment clinical examination.

Delivering stepwise decision making support

The adoption of any CDSSs for infection management highly depends on how well it supports and integrates within the corresponding workflow followed by clinicians to treat infections. The reported infection management pathway follows a stepwise Bayesian like approach [21]. As such, providing comprehensive and specific decision support at each stage of the process will not only facilitate innovation adoption but also influence behaviour change to a greater extent.

Unveiling local AMR patterns to revamp education and awareness

Despite the use of national guidelines to promote prudent use of antimicrobials [14, 15], the disparity in antimicrobial resistance rates between London and the rest of UK is still considerable [17–20]. This disparity also occurs around the world [44]. Therefore, there is a need to analyse hospital-specific microbiology data to assess the actual antimicrobial

resistance patterns (see Table 3.6) and promote rational use of antimicrobials. Furthermore, this information could be used as supporting evidence to refine and tailor the use of national guidelines accordingly.

Promote continuity and education through point-of-care

A wealth of research has demonstrated the importance of effective communication to promote continuity and consistency of care in health centres [45] and it is extremely important in organizations where a wide variety of multidisciplinary health professionals (e.g. nurses, clinicians and infection specialists) are involved in patient care and decision making. As such, the implemented system facilitates access to the complete medical record of the patient at the point of care (see Table 3.3) to promote knowledge sharing, stimulate discussion and assist in shared decision making. In addition, conducted studies have shown that the quality of the health care perceived by patients is highly dependent on establishing trusting, empathetic and reliable relationships [46]. Therefore, the accessibility at point of care also encourages patient participation in decision making (see Table 3.5) and improves their understanding of antimicrobial use, adherence to treatment and self-management.

3.4.1 Providing clinical training to facilitate adoption

The implementation and integration of new technology such as CDSSs is not a one-off event. In fact, it is an ongoing process, which requires support at the organisational level to be effective [47]. As such, EPiC IMPOC has evolved consistently by incorporating new components to tackle the identified weaknesses. Barriers for CDSS uptake are diverse [48]. Most severe barriers include clinicians' perception that the CDSS may reduce their professional autonomy or may be used against them in the event of medical-legal controversies. Less severe are those barrier related with usability problems of the technology interface. With respect to the system usability, EPiC IMPOC might seem complex or confusing on a first sight. In general, this perception is caused by the large amount of clinical evidence and monitoring data presented to clinicians. However, this perceived complexity becomes an advantage once the clinician is familiar with the functionality and meaning of each of the elements within EPiC IMPOC. As such, providing a previous training session to clinicians would greatly benefit innovation adoption and further use of the system in clinical practice.

3.4.2 Limitations

The integration of the system with EHR was tedious. While a large amount of data is being saved in the NHS servers, the access to this information was still very limited. For instance, access to basic vital signs data such as body temperature, respiratory rate or heart rate was not provided forcing clinician to input this information manually leading to an increase in usability time. This led to unpleasant interactions since participants had a limited time on the wards and such information had been previously collected by nurses and other health care workers into the NHS servers. The view of the author is that all the vital signs should be automatically collected and stored in the NHS databases. Moreover, access to that information should be facilitated to clinicians and researchers. In the last year, the NHS commenced the creation of a global API to facilitate recording and access to patient data. Undoubtedly, this API will facilitate the development, implementation and adoption of decision support systems within the NHS. While this

limitation was a constant during the development of the system, specific limitations for each of the modules described previously have been discussed in subsequent chapters.

3.5 CONCLUSION

In summary, EPiC IMPOC has the potential to assist health care providers to practise evidence-based and personalized medicine by providing timely decision support at point of care on a plethora of aspects within the infection management pathway. The process followed to design EPiC IMPOC had several iterations and included a wide range of health care workers; from nurses to infection specialists. The overall feedback provided was good yet some limitations were identified. In particular, participants highlighted the need to reduce the manual collection of data by retrieving such variables automatically from the NHS databases and the need to provide specific support. The provision of training to clinicians is essential to reduce the perceived complexity, encourage appropriate use of the system and therefore promote its adoption. Finally, the modularity of EPiC IMPOC endows the system with high scalability; that is, inclusion of novel decision support modules or further development of the user interface.

BIBLIOGRAPHY

- [1] Nitika Pant Pai, Caroline Vadnais, Claudia Denkinger, Nora Engel, and Madhukar Pai. "Point-of-care testing for infectious diseases: diversity, complexity, and barriers in low-and middle-income countries". *PLoS medicine* 9.9 (2012), e1001306.
- [2] Carliss Young Baldwin and Kim B Clark. *Design rules: The power of modularity*. Vol. 1. MIT press, 2000.
- [3] Belen Diaz-Agudo, Pedro A Gonzalez-Calero, Juan A Recio-Garcia, and Antonio A Sánchez-Ruiz-Granados. "Building CBR systems with jCOLIBRI". *Science of Computer Programming* 69.1-3 (2007), pp. 68–75.
- [4] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, et al. "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.
- [5] Guillaume Lemaître, Fernando Nogueira, and Christos K. Aridas. "Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning". *CoRR* abs/1609.06570 (2016).
- [6] Wes McKinney. "pandas: a Foundational Python Library for Data Analysis and Statistics" ().
- [7] Wes McKinney. "Data Structures for Statistical Computing in Python". *Proceedings of the 9th Python in Science Conference*. Ed. by Stéfan van der Walt and Jarrod Millman. 2010, pp. 51–56.
- [8] J. D. Hunter. "Matplotlib: A 2D graphics environment". *Computing In Science & Engineering* 9.3 (2007), pp. 90–95.
- [9] Michael Waskom, Olga Botvinnik, drewokane, Paul Hobson, David, Yaroslav Halchenko, et al. *seaborn: v0.7.1 (June 2016)*. June 2016.
- [10] Erik Sundvall, Mikael Nyström, Daniel Karlsson, Martin Eneling, Rong Chen, and Håkan Öрман. "Applying representational state transfer (REST) architecture to archetype-based electronic health record systems". *BMC medical informatics and decision making* 13.1 (2013), p. 57.
- [11] Timothy M Rawson, Luke SP Moore, Bernard Hernandez, Enrique Castro-Sanchez, Esmita Charani, Pantelis Georgiou, et al. "Patient engagement with infection management in secondary care: a qualitative investigation of current experiences". *BMJ open* 6.10 (2016), e011040.
- [12] Timothy M Rawson, Luke SP Moore, Enrique Castro-Sanchez, Esmita Charani, Bernard Hernandez, Vivian Alividza, et al. "Development of a patient-centred intervention to improve knowledge and understanding of antibiotic therapy in secondary care". *Antimicrobial Resistance & Infection Control* 7.1 (2018), p. 43.

- [13] P. Herrero, T. M. Rawson, A. Philip, L. S. P. Moore, A. H. Holmes, and P. Georgiou. "Closed-Loop Control for Precision Antimicrobial Delivery: An In Silico Proof-of-Concept". *IEEE Transactions on Biomedical Engineering* 65.10 (Oct. 2018), pp. 2231–2236.
- [14] European Centre for Disease Prevention and Control. *Surveillance of antimicrobial resistance in Europe 2016. Annual report of the European Antimicrobial Resistance Surveillance Network (EARS-Net)*. Stockholm: ECDC; 2017.
- [15] World Health Organization. *Global antimicrobial resistance surveillance system (GLASS) report: early implementation 2016-2017*. Geneva: WHO; 2017. Licence: CC BY-NC-SA 3.0 IGO.
- [16] Luke SP Moore, Rachel Freeman, Mark J Gilchrist, Myriam Gharbi, Eimear T Branigan, Hugo Donaldson, et al. "Homogeneity of antimicrobial policy, yet heterogeneity of antimicrobial resistance: antimicrobial non-susceptibility among clinical isolates from primary, secondary and tertiary care patients in London". *Journal of Antimicrobial Chemotherapy* 69.12 (2014), pp. 3409–3422.
- [17] Public Health England. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR). Annual report 2014*. London: PHE; 2014.
- [18] Public Health England. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR). Annual report 2015*. London: PHE; 2015.
- [19] Public Health England. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR). Annual report 2016*. London: PHE; 2016.
- [20] Public Health England. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR). Annual report 2017*. London: PHE; 2017.
- [21] Timothy Miles Rawson, Esmita Charani, Luke Stephen Prockter Moore, Bernard Hernandez, Enrique Castro-Sánchez, Pau Herrero, et al. "Mapping the decision pathways of acute infection management in secondary care among UK medical physicians: a qualitative study". *BMC medicine* 14.1 (2016), p. 208.
- [22] Yiannis Kyratsis, Raheelah Ahmad, and Alison Holmes. "Technology adoption and implementation in organisations: comparative case studies of 12 English NHS Trusts". *BMJ open* 2.2 (2012), e000872.
- [23] Annette Moxey, Jane Robertson, David Newby, Isla Hains, Margaret Williamson, and Sallie-Anne Pearson. "Computerized clinical decision support for prescribing: provision does not guarantee uptake". *Journal of the American Medical Informatics Association* 17.1 (2010), pp. 25–33.
- [24] Rosy Tsopra, Jean-Philippe Jais, Alain Venot, and Catherine Duclos. "Comparison of two kinds of interface, based on guided navigation or usability principles, for improving the adoption of computerized decision support systems: application to the prescription of antibiotics". *Journal of the American Medical Informatics Association* 21.e1 (2013), e107–e116.
- [25] Viswanath Venkatesh and Fred D Davis. "A model of the antecedents of perceived ease of use: Development and test". *Decision sciences* 27.3 (1996), pp. 451–481.
- [26] Ben Shneiderman, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist, and Nicholas Diakopoulos. *Designing the user interface: strategies for effective human-computer interaction*. Pearson, 2016.

- [27] Jason J Saleem, Mindy E Flanagan, Nancy R Wilck, Jim Demetriades, and Bradley N Doebbeling. "The next-generation electronic health record: perspectives of key leaders from the US Department of Veterans Affairs". *Journal of the American Medical Informatics Association* 20.e1 (2013), e175–e177.
- [28] Michael A Rubin, Kim Bateman, Sharon Donnelly, Gregory J Stoddard, Kurt Stevenson, Reed M Gardner, et al. "Use of a personal digital assistant for managing antibiotic prescribing for outpatient respiratory tract infections in rural communities". *Journal of the American Medical Informatics Association* 13.6 (2006), pp. 627–634.
- [29] Laura J Hoeksema, Alia Bazzi-Asaad, Edwin A Lomotan, Diana E Edmonds, Gabriela Ramírez-Garnica, Richard N Shiffman, et al. "Accuracy of a computerized clinical decision-support system for asthma assessment and management". *Journal of the American Medical Informatics Association* 18.3 (2011), pp. 243–250.
- [30] Jacques Bouaud, Seroussi B, Antoine E.C., Gozy M, Khayat D, and Boisvieux J.F. "Hypertextual navigation operationalizing generic clinical practice guidelines for patient-specific therapeutic decisions." *Proceedings of the AMIA Symposium*. American Medical Informatics Association. 1998, p. 488.
- [31] Justin Starren and Stephen B Johnson. "An object-oriented taxonomy of medical data presentations". *Journal of the American Medical Informatics Association* 7.1 (2000), pp. 1–20.
- [32] Douglas S Bell and RA Greenes. "Evaluation of UltraSTAR: performance of a collaborative structured data entry system." *Proceedings of the Annual Symposium on Computer Application in Medical Care*. American Medical Informatics Association. 1994, p. 216.
- [33] Jan Horsky, Gordon D Schiff, Douglas Johnston, Lauren Mercincavage, Douglas Bell, and Blackford Middleton. "Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions". *Journal of biomedical informatics* 45.6 (2012), pp. 1202–1216.
- [34] Blackford Middleton, Meryl Bloomrosen, Mark A Dente, Bill Hashmat, Ross Koppel, J Marc Overhage, et al. "Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA". *Journal of the American Medical Informatics Association* 20.e1 (2013), e2–e8.
- [35] Jeffery L Belden, Rebecca Grayson, and Janey Barnes. *Defining and testing EMR usability: Principles and proposed methods of EMR usability evaluation and rating*. Tech. rep. Healthcare Information and Management Systems Society (HIMSS), 2009.
- [36] Luke SP More, Bernard Hernandez, Timothy M Rawson, Pau Herrero, Alison H. Holmes, and Pantelis Georgiou. "Intelligent clinical decision support for antimicrobial prescribing in critical care" ().
- [37] Bernard Hernandez, Pau Herrero, Timothy M. Rawson, Luke S. P. Moore, Esmita Charani, Alison H. Holmes, et al. "Data-driven Web-based Intelligent Decision Support System for Infection Management at Point-Of-Care: Case-Based Reasoning Benefits and Limitations". 5 (2017), pp. 119–127.

-
- [38] Cara B Litvin, Steven M Ornstein, Andrea M Wessell, Lynne S Nemeth, and Paul J Nietert. "Use of an electronic health record clinical decision support tool to improve antibiotic prescribing for acute respiratory infections: the ABX-TRIP study". *Journal of general internal medicine* 28.6 (2013), pp. 810–816.
- [39] Mical Paul, Steen Andreassen, Anders D Nielsen, Evelina Tacconelli, Nadja Al-manasreh, Abigail Fraser, et al. "Prediction of bacteremia using TREAT, a computerized decision-support system". *Clinical Infectious Diseases* 42.9 (2006), pp. 1274–1282.
- [40] Athanasios Tsoukalas, Timothy Albertson, and Ilias Tagkopoulos. "From data to optimal decision making: a data-driven, probabilistic machine learning approach to decision support for patients with sepsis". *JMIR medical informatics* 3.1 (2015), e11.
- [41] Niloofar Mohammadzadeh and Reza Safdari. "Patient monitoring in mobile health: opportunities and challenges". *Medical Archives* 68.1 (2014), p. 57.
- [42] Karen A Wager, Marilyn J Schaffner, Bonnie Foulois, Abby Swanson Kazley, Cheryl Parker, and Helena Walo. "Comparison of the quality and timeliness of vital signs data using three different data-entry devices". *CIN: Computers, Informatics, Nursing* 28.4 (2010), pp. 205–212.
- [43] Ashis Jalote-Parmar, Petra Badke-Schaub, Wajid Ali, and Eigil Samset. "Cognitive processes as integrative component for developing expert decision-making systems: A workflow centered framework". *Journal of Biomedical informatics* 43.1 (2010), pp. 60–74.
- [44] Jim O'Neill. "Antimicrobial resistance: tackling a crisis for the health and wealth of nations". *Review on antimicrobial resistance* (2014), pp. 1–16.
- [45] Michael Leonard, Suzanne Graham, and Doug Bonacum. "The human factor: the critical importance of effective teamwork and communication in providing safe care". *BMJ Quality & Safety* 13.suppl 1 (2004), pp. i85–i90.
- [46] Anthony C Berman and Darryl S Chutka. "Assessing effective physician-patient communication skills: "Are you listening to me, doc?"". *Korean journal of medical education* 28.2 (2016), p. 243.
- [47] Alison Porter, Jeremy Dale, Theresa Foster, Pip Logan, Bridget Wells, and Helen Snooks. "Implementation and use of computerised clinical decision support (CCDS) in emergency pre-hospital care: a qualitative study of paramedic views and experience using Strong Structuration Theory". *Implementation Science* 13.1 (2018), p. 91.
- [48] Elisa G Liberati, Francesca Ruggiero, Laura Galuppo, Mara Gorli, Marien González-Lorenzo, Marco Maraldi, et al. "What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation". *Implementation Science* 12.1 (2017), p. 113.

ANTIMICROBIAL RESISTANCE SURVEILLANCE

This chapter describes the research towards the implementation of a decision support module that performs antimicrobial resistance (AMR) surveillance from susceptibility test data. Firstly, the chapter describes the most widespread methods used to measure AMR (section 4.1). Then, it presents a number of strategies to enhance AMR surveillance such as quantification of antimicrobial spectrum of activity (section 4.5) or estimation of resistance trends through regression analysis (section 4.6). To conclude, the chapter validates the outcomes obtained using the susceptibility test data provided by the Imperial College Healthcare National Health Service Trust in London with those reported by renown health care organizations and independent research articles within the literature.

4.1 HOW TO MEASURE AMR?

The growing threat of antimicrobial resistance (AMR) is a leading patient health and safety issue, with estimates that AMR will be responsible for more than 10 million deaths by 2050 [1]. A major driver of AMR has been the misuse of antimicrobials in humans [2]. Whilst reasons for the misuse of antimicrobials are complex and multifaceted, a number of factors have been described and investigated. At the individual level, physicians often prioritise the management of the patient being treated, paying little regard to the long-term consequences of overusing antimicrobials [3]. Moreover, the majority of antimicrobials are prescribed by individuals who are not experts in infection management and may have limited understanding of antimicrobials and the potential consequences of AMR [2, 4–6]. At the hospital level, a number of barriers to the effective use of antimicrobials have been described, including the role of team hierarchies, functional communication and prescribing etiquette [7–9]. To address the challenges posed by AMR, the importance of behaviour change interventions to improve the long-term use of antimicrobials in infection management has been recognised [8–10]. For such reason, numerous health organizations have promoted antimicrobial surveillance to regulate prescriptions within clinical practice. At national level, Public Health England implemented the English surveillance program for antimicrobial utilisation and resistance (ESPAUR) which provides annual reports as a benchmark to determine appropriate local action [11–14]. At international level, the European Centre for Disease Prevention and Control through the European antimicrobial resistance surveillance network (EARS-Net) has created the largest publicly funded system for antimicrobial surveillance in Europe [15–18]. Furthermore, the World Health Organization has recently implemented the global anti-

crobial resistance surveillance system (GLASS) [19] to strengthen the evidence base on AMR and inform decision-making.

With increasing electronic recording of data, there is a growing interest in the potential secondary use of microbiology records to provide the necessary information to support antimicrobial stewardship programs [20]. These programs are crucial to guide health care organizations designing evidence-based policies to combat AMR [21, 22]. In particular, susceptibility reporting has shown to be a determinant data source to inform empiric antimicrobial therapy selection [23]. The most widespread resistance measurement is denoted as single antimicrobial resistance index (SARI) and evaluates the proportion (or percentage) of hosts harbouring resistant pathogens within a certain population (see left column in Table 4.1). In scenarios where a pathogen exhibits resistance to numerous antimicrobials, the multiple antimicrobial resistance index is used to evaluate the ratio of antimicrobials to which a pathogen is resistant [24] (see middle column in Table 4.1). Moreover, the drug resistance index (DRI) aims to measure the proportion of resistant pathogens by incorporating antimicrobial use information into the equation [25] (see right column in Table 4.1). These metrics inform clinicians on overall antimicrobial resistance levels; however, they overlook information such as resistance tendency or seasonality. For such purpose, the previously explained indexes are computed on consecutive and independent time intervals to produce resistance time series signals. Unfortunately, these are predominantly analysed by means of visual graphs with limited exploitation of computational algorithms to automate handling of large datasets [1, 26, 27].

Table 4.1: Description of indexes to measure AMR.

Single Antibiotic Resistance	Multiple Antibiotic Resistance	Drug Resistance
This index describes the proportion of resistant isolates for a given set of susceptibility tests. It provides a value within the range [0,1] where values close to one indicate high resistance. It is agnostic to pathogen, antibiotic and time. The variables R, I and S represent the number of susceptibility tests with Resistant, Intermediate and Susceptible outcomes respectively. The definition might vary slightly since the intermediate category is not always considered.	This index describes the ratio of antimicrobials tested (T) to which a pathogen is resistant (R). It provides a value within the range [0,1] where values close to one indicate high multi-drug resistance. It highly depends on the antimicrobials to which the pathogen is tested. Since tested antimicrobials vary among health care centres and time, comparison and analysis of its evolution in time is not straight forward. In addition, antibiotics which are intrinsically resistant should not be considered.	This index measures the proportion of pathogens that are resistant to the antimicrobials used to treat them. It provides a value within the range [0,1] where values close to one indicate high resistant for frequent antimicrobials. The variable ρ_{ik} is the proportion of resistance among organism i to antimicrobial k and q_{ik} is the frequency of drug k used to treat organism i .
$SARI = \frac{R + I}{R + I + S}$	$MARI \Big _{ISO} = \frac{R}{T}$	$DRI \Big _{PAT} = \sum_k \rho_{ik} q_{ik}$

Keys: I=intermediate; R=resistant; S=sensitive; ISO=isolate; T=total; PAT=pathogen;

4.2 MATERIALS AND METHODS

4.2.1 Susceptibility test data

Susceptibility test records (see Figure 4.1) are composed by laboratory identification number (LID), patient identification number (PID), date, sample type or culture (e.g. blood or urine), pathogen, antimicrobial, reported status and outcome (resistant, sensitive or intermediate). In this research, the susceptibility test data were grouped firstly by sample type. Moreover, for each sample type, the data were grouped by pairs (pathogen-antimicrobial) since it is widely accepted by clinicians as detailed in the UK five year strategy in AMR [21].

The data was provided by the Imperial College Healthcare NHS Trust, which comprises three separate hospitals, totalling 1500 beds and serving a population of 2.5 million citizens. Laboratory operating procedures followed national standards for microbiological investigation [28]; isolates were identified using API® (bioMérieux) from 2009 to 2011 and by MALDI-TOF spectroscopy (Biotyper®, Brunker) from 2011 to 2015. The susceptibilities were determined by disc diffusion using BSAC criteria [29]. Results were de-duplicated for organisms repeatedly isolated from the same patient. The automated analysis presented in further sections was performed on more than 3.5 million susceptibility tests for over 300,000 isolates corresponding to approximately 200,000 individuals.

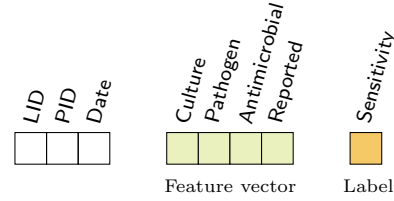


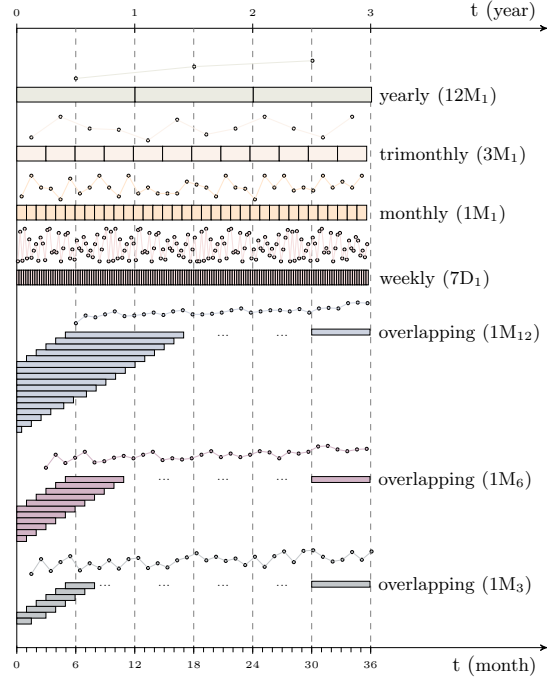
Figure 4.1: Susceptibility test record. It is composed by metadata, sample type, pathogen, antimicrobial and sensitivity outcome (resistant, intermediate or sensitive). *LID*=laboratory identification number; *PID*=patient identification number.

4.2.2 Generation of resistance time series signals

In order to study the temporal evolution of AMR, it is necessary to generate a resistance time series from the susceptibility test data. This is often achieved by computing the resistance index on consecutive partitions of the data. Note that each partition contains the susceptibility tests required to compute a resistance index. The traditional strategy of dealing with partitions considers independent time intervals (see yearly, monthly or weekly time series in Table 4.2). Unfortunately, this strategy forces to trade-off between granularity (level of detail) and accuracy. On one side, weekly time series are highly granular but inaccurate. On the other hand, yearly time series are accurate but rough. Note that the granularity is represented by the number of observations in a time series while the accuracy is closely related with the number of susceptibility tests used to compute the resistance index. Conversely, the overlapping time intervals strategy drops such dependence by defining a window of fixed size which is moved across time. The length of the window is denoted as *period* and the time step as *shift*. For instance, three time series obtained using the overlapping time intervals strategy with a monthly shift (1M) and window lengths of 12, 6 and 3 have been presented for the sake of clarity (see 1M₁₂, 1M₆ and 1M₃ in Table 4.2).

Table 4.2: Description of strategies to generate resistance time series.

Independent time intervals			
This is the traditional method used in antimicrobial surveillance systems where the time spans considered are independent; that is, they do not overlap (e.g. month or year).			
Overlapping time intervals			
This method is defined as a fixed region which is moved across time to compute consecutive resistance indexes. It is described by two parameters; the length of the region (period) and the distance between consecutive windows (shift).			
Notation			
The notation to define the time series generation methodology use is $\text{shift}_{\text{period}}$. Some examples are presented below.			
$\text{shift}_{\text{period}}$	Shift	Period	Type
$1D_1$	Daily	previous day	I
$7D_1$	Weekly	previous week	I
$1M_1$	Monthly	previous month	I
$12M_1$	Yearly	previous year	I
$1M_{12}$	Monthly	previous year	O
$1M_6$	Monthly	previous 6 months	O
$1M_3$	Monthly	previous 3 months	O



Keys: D=day; M=month; I=independent time intervals; O=overlapping time intervals;

The notation to define the time series generation methodology ($\text{shift}_{\text{period}}$) is described with various examples in Table 4.2. For instance, $7D_4$ defines a time series with weekly resistance indexes (7D) calculated using the microbiology records available for the previous four weeks ($4 \times 7D$). It is important to note that some notations are equivalent representations of the same susceptibility data at different granularities, hence their trends are comparable. As an example, the trend estimated for $1M_1$ should be approximately thirty times the trend estimated for $1D_{30}$.

4.2.3 Trend and stationarity in time series

An analysis of stationarity around a trend was carried out to identify time series satisfying the assumptions posed by ARIMA. The augmented Dickey–Fuller test (ADF) was used to determine the presence of a unit root. When the other roots of the characteristic function lie inside the unit circle the first difference of the process is stationary. Due to this property, these are also called difference-stationary processes. Since the absence of unit root is not a proof of non-stationarity, the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test was used to identify the existence of an underlying trend which can also be removed to obtain a stationary process. These are called trend-stationary processes. In both, unit-root and trend-stationary processes, the mean can be increasing or decreasing over time; however, in the presence of a shock, trend-stationary processes revert to this mean tendency in the long run (deterministic trend) while unit-root processes have a permanent impact (stochastic trend). The significance level of the tests was set to 0.05.

4.2.4 Pearson correlation coefficient

It measures the linear correlation between two variables with a value within the range $[-1, 1]$. Coefficient values of -1 , 0 and 1 indicate total negative linear correlation, no linear correlation and total positive correlation respectively. In this study, the coefficient is used to assess whether or not there is a linear correlation between the number of observations (susceptibility test records) and the computed resistance index.

4.2.5 Sigmoid function

The sigmoid function defines those mathematical functions with a characteristic 'S'-shaped curve (see Figure 4.2). However, the term sigmoid function often refers to the special case of the logistic function as stated in Equation 4.1 where α represents maximum value, β determines the steepness and γ the value in the x-axis of the sigmoid midpoint. Although the domain of the sigmoid function contains all real numbers, most often the return value is monotonically increasing in $[0, 1]$ (or alternatively $[-1, 1]$).

$$f(x) = \frac{\alpha}{1 + e^{-\beta(x-\gamma)}} \quad (4.1)$$

The confidence of a resistance index relies on the number of susceptibility test records available. For such reason, the sigmoid function has been used to define the contribution of resistance indexes proportionally to their population size. Lets consider two vectors; one with the resistance indexes ($\vec{r}$) and the other with the corresponding number of susceptibility tests ($\vec{f}$). The implementation of the sigmoid function defines the steepness based on the percentiles of $\vec{f}$ (see Figure 4.2). The values of the lower and upper percentile indicate where the curve should begin to flatten. Similarly, a threshold parameter based on percentiles has been implemented. In this case, the values of the lower and upper percentile define the region in which the sigmoid function is computed. Outside this region either the minimum or maximum weights are returned. This function is used to define the weights in the weight least-squares regression further described in this chapter. In this way, the resistance indexes with extremely low number of observations are discarded. On the contrary, reliable resistance indexes are assigned the maximum weight regardless of the number of observations.

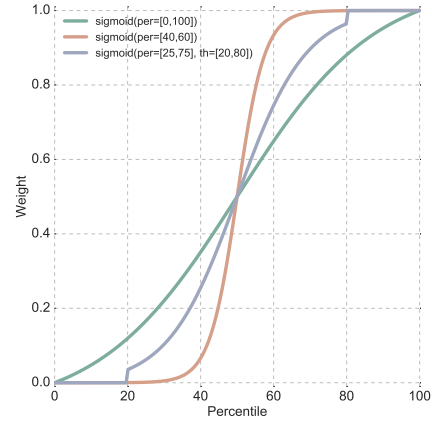


Figure 4.2: Sigmoid function. The sigmoid examples are from top to bottom: (i) percentiles $[0,100]$ (ii) percentiles $[40,60]$ and (iii) percentiles $[25,75]$ with thresholds $[20,80]$. The scale ($\alpha=1$) and shift ($\gamma=0.5$) remained constant.

4.2.6 Statistical significance among regression methods

The statistical significance between the different scenarios was determined using the non parametric test Wilcoxon-Mann-Whitney (also denoted Mann-Whitney U) were the significance level was set to 0.05 to assess the robustness of the evaluated regression analysis methods.

4.2.7 Outline of the experiments conducted

This section briefly outlines the experiments further described in this chapter. For such purpose, the experiments have been depicted graphically in Figure 4.3. The first experiment provides an insight into the susceptibility test data provided by the Imperial College NHS Trust. After this, the second experiment (Experiment II in figure 4.3) computes and display the antimicrobial resistance rates. As such, the susceptibility test data has been grouped in pairs pathogen-antimicrobial and the proportion of resistant isolates have been computed. For the sake of clarity, an example describing the antimicrobial resistance rates for those pairs tested in urine cultures has been presented (see the clustered heatmap in Figure 4.5). The third experiment (Experiment III in Figure 4.3) uses those previously computed resistance rates to measure the antimicrobial spectrum of activity which refers to the range of microbe species that are susceptible to these agents. As an example, the spectrum of activity for those antimicrobials tested within urine cultures are presented (see Table 4.7). Note that the existing spectrum of activity uses the ambiguous narrow-broad categorisation.

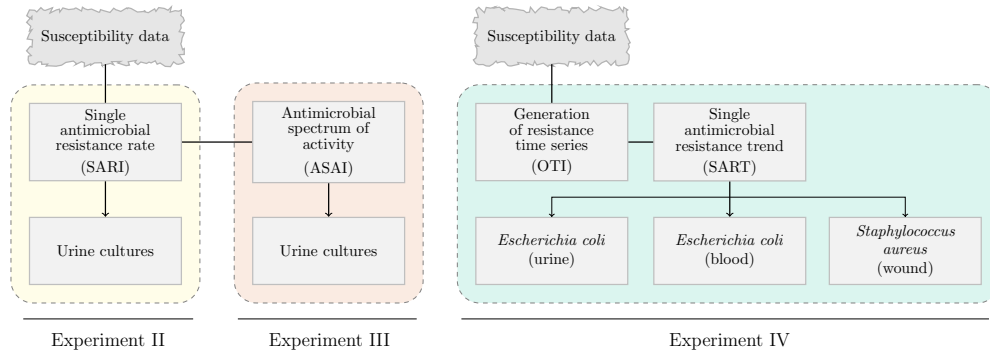


Figure 4.3: Graphical description of the experiments. The experiments described in this diagram are: (i) computation of antimicrobial resistance rates (ii) computation of antimicrobial spectrum of activity and (iii) estimation of antimicrobial resistance trends. The experiments use susceptibility test data. SARI=single antimicrobial resistance index; ASAI=antimicrobial spectrum of activity index; OTI=overlapping time intervals; SART=single antimicrobial resistance trend;

To conclude, the last experiment (Experiment IV in Figure 4.3) aims to estimate the antimicrobial resistance trend for individual pairs. As such, the susceptibility test data is grouped into time intervals to generate a resistance time series. Then, the robustness of various regression analysis techniques to estimate the antimicrobial resistance trend is compared. To demonstrate the validity of the results three different scenarios have been considered: *Escherichia coli* in urine cultures, *Escherichia coli* in blood cultures and *Staphylococcus aureus* in wound cultures. The results obtained for these scenarios are described and compared with those reported by national and international health care organizations.

4.3 EXPERIMENT I: OVERVIEW OF SUSCEPTIBILITY TEST DATA

This section provides an insight into the susceptibility test data provided by the Imperial College London NHS Trust. The main variables within a susceptibility test record are sample type, pathogen, antimicrobial and sensitivity outcome (see Figure 4.1). The analysis was performed on more than 3.5 million susceptibility tests for over 300,000 isolates corresponding to approximately 200,000 individuals.

A total of 3,510,472 susceptibility test records were included in the dataset. From these records, the most commonly requested sample types were urine (1,051,550; 30%), wound (911,297; 26%), sputum (244,348; 7%) and blood (220,470; 6%). The remaining susceptibility test records (1,082,807; 31%) correspond to 114 different sample types. Note that specimen collection is conducted on a wide variety of sites which are closely related with the source of the infection.

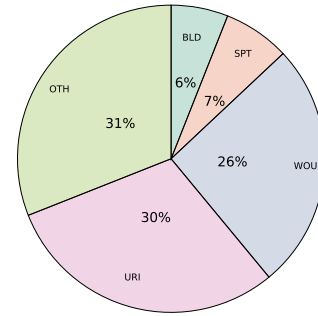


Figure 4.4: Proportion of sample types (cultures). The proportion of cultures within the susceptibility test data. BLD=blood; SPT=sputum; WOU=wound; URI=urine; OTH=other.

The majority of susceptibility tests belong to a reduced set of pathogens which are a common cause of infection in the hospital. For the sake of clarity, the proportion of records for the 10 most common pathogens have been presented in Tables 4.3 (urine), 4.4 (wound), 4.5 (sputum) and 4.6 (blood). For instance, in urine cultures *Escherichia coli* (56%) is the most representative pathogen with more than half of records. A similar result is depicted in wound cultures with *Staphylococcus aureus* (51%). On the other hand, the proportions in both sputum and blood cultures are relatively balanced with *Pseudomonas aeruginosa* (17%) and coagulase negative staphylococcus (28%) as top pathogens respectively. It is surprising how the top 10 more commonly tested pathogens hold the 90%, 81%, 75% and 79% of susceptibility test records in urine, wound, sputum and blood respectively. Note that some of the codes presented in the tables below correspond to the genus (*KLEBS*=*Klebsiella*) though the majority also specify the species (*KPNE*=*Klebsiella pneumoniae*). The codes used to represent organisms and antimicrobials are described in Appendix D.

Table 4.3: Urine cultures

Organism	Records
ECOL	586280 (56%)
ENTC	115926 (11%)
COLIF	105706 (10%)
PAER	26044 (2%)
BHSB	24572 (2%)
KPNE	20620 (2%)
PROTE	18544 (2%)
PMIR	18405 (2%)
CNS	18194 (2%)
SAUR	15216 (1%)

Description: The number of records (percentage) for top 10 pathogens in urine.

Table 4.4: Wound

Organism	Records
SAUR	464492 (51%)
PAER	77241 (8%)
COLIF	41700 (5%)
ECOL	38082 (4%)
NGON	25983 (3%)
LFC	19600 (2%)
BHSB	18761 (2%)
PSEUD	16501 (2%)
ENTC	16477 (2%)
ENTB	16476 (2%)

Description: The number of records (percentage) for top 10 pathogens in wound.

Table 4.5: Sputum

Organism	Records
PAER	42362 (17%)
SAUR	38928 (16%)
ECOL	24059 (10%)
HINF	17863 (7%)
KPNE	17723 (7%)
COLIF	12924 (5%)
KLEBS	8767 (4%)
SMAR	7895 (3%)
SPNE	7444 (3%)
ENTB	7286 (3%)

Description: The number of records (percentage) for top 10 pathogens in sputum.

Table 4.6: Blood cultures

Organism	Records
CNS	60814 (28%)
ECOL	59315 (27%)
SAUR	20383 (9%)
KPNE	10298 (5%)
KLEBS	6697 (3%)
PAER	5186 (2%)
ENTC	3355 (2%)
MICROC	2998 (1%)
ENTB	2892 (1%)
AHS	2697 (1%)

Description: The number of records (percentage) for top 10 pathogens in blood.

4.4 EXPERIMENT II: ANTIMICROBIAL RESISTANCE RATES

This experiment computes the antimicrobial resistance rates (proportion of resistance isolates as defined by SARI in Table 4.1) for specific pathogen-antimicrobial pairs. Note that the resistance rates were computed independently for each sample type (e.g. urine or blood). A graphical representation of the resistance rates for pairs tested in urine samples is shown in Figure 4.5. The x-axis represents antibiotics (e.g. imipenem) and the corresponding category (e.g. carbapenems). The y-axis represents the genus of the pathogen (e.g. pseudomonas) and the species (e.g. aeruginosa). The magnitude of the resistance is described with different shades of red; the darker the shade the higher the resistance. Each row describes the resistance of the pathogen to various antibiotics (pathogen resistance profile). Similarly, each column describes the effectiveness of the antibiotic against various pathogens (antibiotic effectiveness profile).

The pathogens (y-axis) and antibiotics (x-axis) presented in Figure 4.5 have been clustered based on the similarity of the profiles. First, the similarity between either pathogen resistance profiles (rows) or antimicrobial effectiveness profiles (columns) was computed using the braycurtis distance. Then, the Weighted Pair Group Method with Arithmetic Mean (WPGMA) agglomerative hierarchical clustering was applied. The arrangement of the clusters produced is represented using dendograms (tree diagrams). In addition, colour labels have been included to represent the clusters. For pathogens, the clusters are the genera (e.g. klebsiella, pseudomonas or enterococcus). For the antibiotics, the clusters are the categories (e.g. penicillins, carbapenems or cephalosporins).

A number of patterns within susceptibility testing arise from the results presented in Figure 4.5. Firstly, two main areas are visible within the antimicrobial resistance heat map (red shades). The main area (largest square) is formed from a wide variety of pathogens (around 18 pathogens belonging to 10 different species) which are consistently tested against the same group of antimicrobials. From these pathogens, approximately half (two citrobacters, three enterobacters, two serratia and one morganelia) are resistant to amoxicillin-clavunilate (AUG), cephalixin (CELX) and cefoxitin (CXT). Note that ESBL are enzymes that mediate resistance to extended spectrum cephalosporins and monobactams. Moreover, klebsiella (in particular *Klebsiella pneumoniae* – KPNE) and *Acinetobacter baumannii* (ABAU) are resistant to the majority of the tested antimicrobials. The second area (rectangle on the top right) is formed by *Staphylococcus saprophyticus* (SSAP), coagulase negative staphylococcus (CNS) and *Staphylococcus aureus* (SAUR). These three pathogens are tested against the same set of antimicrobials and present very similar resistance profiles. For instance, they are all highly resistant to penicillin and to some extent to erythromycin. There are also spare pathogens and antimicrobials without clear specific patterns.

To summarise, the antimicrobials selected for susceptibility testing highly depend on the infectious pathogen. The hierarchical clustering method applied groups pathogens with similar resistance profiles together. Moreover, the similarity between two pathogens is boosted when the same antimicrobials are tested. This is reflected by the appropriate grouping of the cluster labels representing the genera of pathogens.

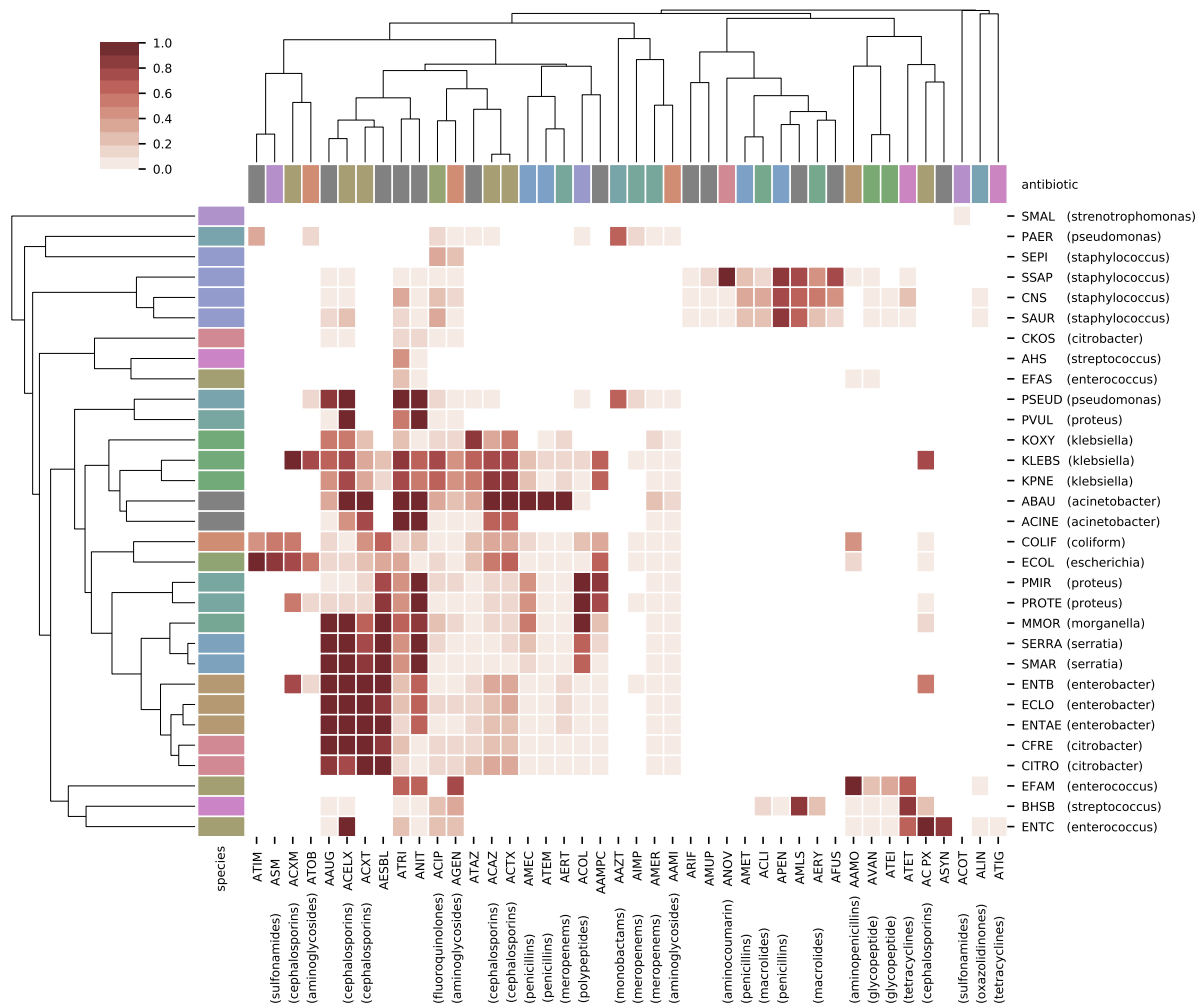


Figure 4.5: Antimicrobial resistance rates in urine cultures (clustermap). The x-axis represents antibiotics (e.g. imipenem) and the corresponding category (e.g. carbapenems). The y-axis represents the genus of the pathogen (e.g pseudomonas) and the species (e.g. aeruginosa). The magnitude of the resistance is described with different shades of red; the darker the shade the higher the resistance. Exclusively those resistance rates calculated with at least a hundred isolates have been displayed. Pathogens and antimicrobials were clustered using the Weighted Pair Group Method with Arithmetic Mean (WPGMA) agglomerative hierarchical clustering. The distance metric used was braycurtis. The arrangement of the clusters produced through hierarchical clustering is represented using dendograms (tree diagrams). Colour labels have been included to facilitate visual identification of clusters. For pathogens, the clusters represent genera (e.g. klebsiella, pseudomonas or enterococcus). For the antibiotics, the clusters represent categories (e.g. penicillins, carbapenems or cephalosporins).

4.5 EXPERIMENT III: ANTIMICROBIAL SPECTRUM OF ACTIVITY

The antimicrobial spectrum of activity refers to the range of microbe species that are susceptible to these agents and therefore can be treated. In general, antimicrobial agents are classified into broad, intermediate or narrow spectrum. Broad spectrum antimicrobials are active against both Gram-positive and Gram-negative bacteria. In contrast, narrow spectrum antimicrobials have limited activity and are effective only against particular species of bacteria. While these profiles appeared in the mid-1950s, little effort has been made to define them. Furthermore, such ambiguous labels are overused for different and even contradictory purposes [30].

Recently, the antibiotic spectrum index (ASI) was defined to measure the spectrum of activity of most common pathogens [31]. For such purpose, they assigned a point value when the antimicrobial was effective to treat 12 conditions. The effectiveness was assessed by infectious specialists and discrepancies were resolved through review of the published data and pharmacy databases. The conditions varied in specificity such as (i) treat enterococcus (ii) treat methicillin resistant *Staphylococcus aureus* or (ii) treat either *Enterobacter serriatia* or citrobacter. In addition, two extra points were added for atypical and multi-drug resistance. As a result the ASI score was an integer value within the range [0,14]; that is, 15 possible categories. This categorization extends the widespread narrow-broad classification, yet it is incomplete. First, it only considers a small subset of common bacteria. Second, the values are still discrete (15 possible categories). Also, the definition of effectiveness based on experts knowledge was vague and sub-optimal. As such, the antimicrobial spectrum of activity index (ASAI) has been defined to quantify the spectrum of activity automatically from susceptibility test data.

4.5.1 The antimicrobial spectrum of activity index

In order to compute the antimicrobial spectrum of activity index (ASAI), it is necessary to previously obtain the overall resistance (SARI) for all the microbe-antimicrobial pairs. Furthermore, by following the criteria used in the narrow-broad approach, these pairs were grouped into Gram-positive and Gram-negative. The methodology to compute the ASAI (see Figure 4.6) was applied to each category independently. Briefly, the weighted proportion of species to which the antimicrobial is effective is computed for each genus. These are later added up and normalized by the number of genera tested. An antimicrobial is considered effective to treat a particular species when the corresponding resistance index (SARI) is lower than a given threshold.

This process is defined mathematically in Equation 4.2 where th is the effective threshold, N_{gn} is the total number of genera, N_{sp}^{gn} is the total number of species for a particular genus, w_{gn} is the weight for the genus and w_{gn}^{sp} is the weight of a species within a genus. The weight parameters remained uniform for the examples presented in this section. Note that pairs with inherent resistance, and therefore a SARI of 1, are not considered. The ASAI ranges within the interval [0,1].

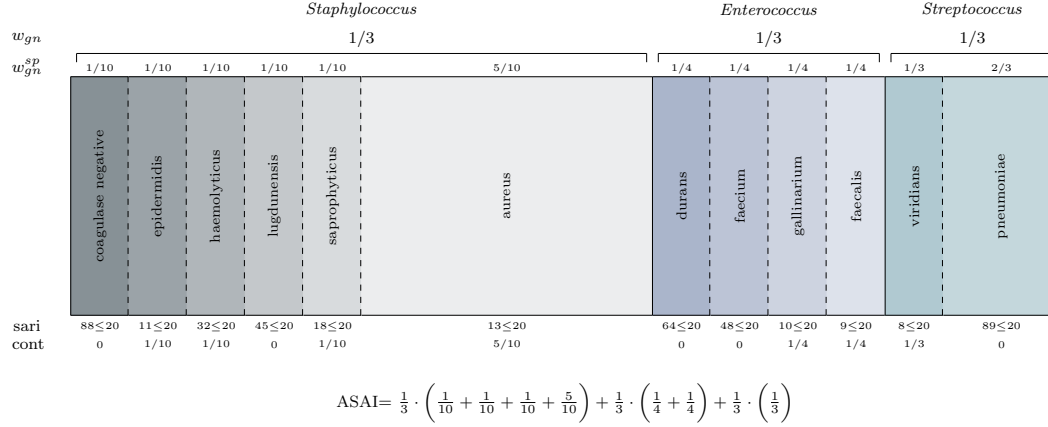


Figure 4.6: Antimicrobial spectrum of activity: graphical description. The illustration presents three different genus (*Staphylococcus*, *Enterococcus* and *Streptococcus*) with a number of species. Genera are represented by colours and species through different intensities. The single antimicrobial resistance index (SARI), the corresponding partial contribution (cont) and the antimicrobial spectrum of activity index (ASAI) are also unveiled. An antimicrobial was considered effective when the resistance (SARI) was below 20%.

$$ASAI_{th} = \frac{1}{N_{gn}} \sum_{\forall gn} w_{gn} \cdot \left(\frac{1}{N_{sp}^{gn}} \sum_{\forall sp^{gn}} w_{sp}^{gn} \cdot \text{effective}_{th}(SARI_{sp}^{gn}) \right)$$

where

$$\sum_{\forall gn} w_{gn} = 1 \quad \sum_{\forall sp} w_{sp}^{gn} = 1 \quad \text{effective}_{th}(SARI) = \begin{cases} 1 & \text{if } SARI \leq th \\ 0 & \text{otherwise} \end{cases} \quad (4.2)$$

4.5.2 Spectrum of activity summary for urine culture

The antimicrobial spectrum of activity has been presented for urine cultures since it was the most frequent culture (30%) within the provided NHS data (see Table 4.7). The table contains for each antimicrobial the Gram-negative (ASAI_N) and Gram-positive (ASAI_P) spectrum indexes. In addition, the antimicrobial categories in which they were classified within the literature (see columns narrow, intermediate and broad) and the overall spectrum category are presented. A graphical summary has been included to facilitate visualization (see graph on the right in Table 4.7). The geometric mean of ASAI_N and ASAI_P was used to sort antimicrobials by combined spectrum of activity.

The data considered in this study was not collected purposely. As such, antimicrobials are often evaluated on those microbes to which they are known to be effective. This behaviour might lead to bias. For such reason, the number of genera (N_{gn}) and species (N_{sp}) to which an antimicrobial has been tested are presented for clarity. For instance, widely accepted narrow-spectrum antibiotics such as vancomycin and erythromycin were tested on just one Gram-negative species (*Escherichia coli*). Moreover, aztreonam was not tested on Gram-positive microbes. These irregularities have not been included in the graphical summary (see † in Table 4.7).

Table 4.7: Antimicrobial spectrum of activity (ASAI) summary in urine cultures.

Antimicrobial	ASAI _N ($N_{gn}-N_{sp}$)	ASAI _P ($N_{gn}-N_{sp}$)	Narrow	Intermediate	Broad	Category	ASAI _N	ASAI _P
Tazocin (TAZ)	71.4 (7–22)	100.0 (3–4)			[32, 33]	Broad		
Meropenem (MER)	83.9 (7–22)	75.0 (4–5)			[32, 33]	Broad		
Amikacin (AMI)	75.0 (8–23)	66.7 (3–4)				Broad		
Ceftazidime (CAZ)	64.0 (7–22)	66.7 (3–4)			[32]	Broad		
Ertapenem (ERT)	56.7 (6–18)	75.0 (4–5)				Broad		
Amoxicillin (AMO)	50.0 (4–5)	73.8 (5–27)		[34]	[32]	Intermediate		
Gentamicin (GEN)	63.7 (9–24)	54.0 (6–27)		[33]		Intermediate		
Tigecycline (TIG)	33.3 (3–5)	100.0 (2–4)				Intermediate		
Amoxicillin-Clavunilate (AUG)	38.1 (7–18)	86.3 (6–31)		[34]	[33]	Intermediate		
Nitrofurantoin (NIT)	36.2 (7–17)	66.7 (6–32)		[33]		Intermediate		
Cefotaxime (CTX)	24.7 (5–18)	75.0 (4–5)				Intermediate		
Cephalexin (CELX)	33.3 (6–14)	54.2 (6–29)				Intermediate		
Ciprofloxacin (CIP)	34.9 (9–25)	27.0 (6–26)			[32, 35]	Intermediate		
Mecillinam (MEC)	17.5 (6–18)	50.0 (2–3)				Intermediate		
Cefoxitin (CXT)	11.7 (5–16)	68.6 (4–8)			[34]	Intermediate		
Trimethoprim (TRI)	20.1 (8–20)	13.8 (5–29)		[33]		Intermediate		
Vancomycin (VAN)	100.0 (1–1) [†]	75.0 (5–25)	[33, 34]			Narrow		
Aztreonam (AZT)	41.5 (7–10)	–	[34]		[32]	Narrow		
Clindamycin (CLI)	100.0 (1–1) [†]	36.0 (5–27)			[33]	Narrow		
Erythromycin (ERY)	100.0 (1–1) [†]	34.0 (5–26)				Narrow		

[†] represents insufficient number of species not displayed in graphical summary.

Note: The effective threshold was set at 0.05.

Note: The antimicrobials have been sorted using the geometric mean of the indexes.

Keys: ASAI_N=antimicrobial spectrum of activity index for Gram-negative; ASAI_P=antimicrobial spectrum of activity index for Gram-positive; N_{gn} =number of genera; N_{sp} =number of species;

In overall, the computed antimicrobial spectrum of activity indexes align with those reported in the literature (see Table 4.7). Note that the antimicrobial effectiveness (see Equation 4.2) was assessed based on the hospital-specific resistance rates previously computed. Thus, the antimicrobial spectrum of activity is also tailored to the local resistance patterns. In addition, it provides deeper understanding of the antimicrobial coverage for the different groups under analysis. For instance, the graph in Table 4.7 displays the effectiveness of the antimicrobials for the two groups selected: Gram-positive and Gram-negative. This approach can be applied to any number of groups.

4.5.3 Enhancing traditional antimicrobial spectrum categorization

Empirical therapies commonly rely on broad spectrum antimicrobials to provide treatment active to most likely microbes. As such, these therapies need to cover all potential resistant pathogens. Later, these broad-spectrum therapies are streamlined (i.e. de-scaled) on the basis of microbiological data and clinical response. However, the existing categorization based on the terms broad and narrow applied to the spectrum of antimicrobials has been reported to be ambiguous and sometimes misleading and therefore needs revision [30, 36]. Antimicrobials are often assigned to a different category depending on the study (see Table 4.7). This ambivalence mostly occurs when an antimicrobial lies within contiguous categories. For instance, incertitude within the intermediate-broad region (amoxicillin, trimethoprim or amoxicillin-clavulanate) is frequent. This issue also occurs in the narrow-intermediate region (trimethoprim). To resolve such ambiguity, the ASAI provides a continuous numerical value within the range [0,1]. On the other hand, the study domain might also produce the aforementioned ambiguity. For instance, aztreonam which is commonly considered a narrow spectrum antimicrobial was assigned to the broad category in [32]. Since the focus of the previous study was to rank the activity of 10 broad spectrum antimicrobials against Gram-

negative bacteria, such miscategorization from narrow to broad is justifiable. Note that in the results (see Table 4.7) the activity of aztrenonam against Gram-negative bacteria (41.5) is higher than many other antimicrobials. In order to address such ambiguity, the problem domain has been divided into two sub-domains (Gram-positive and Gram-negative) yet any number of domains can be selected.

To conclude, it is essential to emphasize that the choice of antimicrobial should be guided by the targeted organs [30] and the corresponding local or national resistance surveillance data. This is crucial since the disparate levels of resistance around the world [1] directly affect the spectrum of activity of the antimicrobials.

4.6 EXPERIMENT IV: ANTIMICROBIAL RESISTANCE TRENDS

The diagram in Figure 4.7 describes the methodology implemented to estimate secular trends in AMR from susceptibility data. Since data corruption might occur in clinical environments, those susceptibility test records wrongly reported (human or device errors) or duplicated were discarded. The remaining records were divided into combinations (tuples defined by sample type, pathogen and antimicrobial) for which a resistance time series signal was generated using either independent or overlapping time intervals (see subsection 4.2.2). The time series were linearly interpolated to fill sporadic missing values. No additional filters or preprocessing steps were applied. An analysis of stationarity around a trend was carried out to identify interesting combinations and regression analysis was used to quantify its tendency.

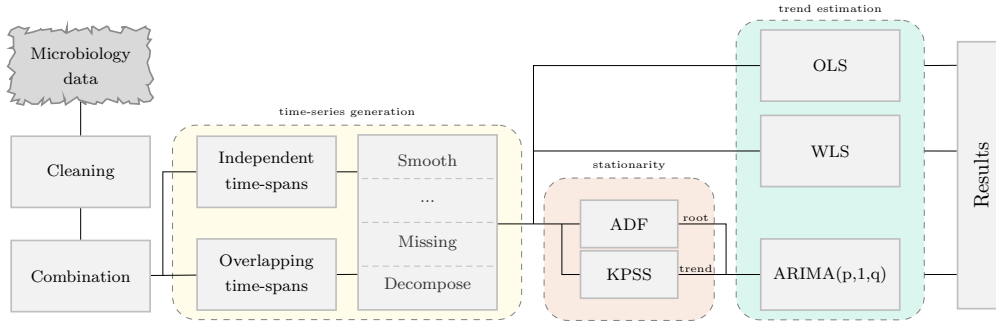


Figure 4.7: High-level methodology diagram for antimicrobial resistance trend estimation. It is composed by three main sections: time series generation (yellow), stationarity analysis (orange) and trend estimation (green). The stationarity analysis was performed using the Augmented Dickey-Fuller (ADF) and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests to identify root-stationary and trend-stationary time series signals. The regression analysis methods considered were ordinary least squares (OLS), weighted least squares (WLS) and autoregressive integrated moving average (ARIMA).

4.6.1 Regression analysis for trend estimation

The linear model (see Equation 4.3) has been selected to quantify resistance tendency for several reasons: (i) the development of resistance in pathogens is an evolutionary response to the selective pressure of antimicrobials, hence large variations in short periods (e.g. consecutive days or months) are not expected (ii) the slope parameter can be directly translated to change over time increasing its practicability and (iii) the offset parameter is highly related with the overall resistance. Hence, the response variable in regression analysis (resistance index) is described by the explanatory variable (time). The slope (m) ranges within the interval $[-1,1]$ where sign and absolute value capture direction and rate of change respectively. The unit of the slope is represented by Δ_y/Δ_x . It has been denoted as Single Antimicrobial Resistance Trend (SART).

$$y = mx + n \quad \text{where} \quad m = \frac{y_{t+1} - y_t}{x_{t+1} - x_t} \quad (4.3)$$

Least Squares Regression

The optimization problem in ordinary least squares (OLS) regression minimizes the least square errors to find the best fitting model as described in Equation 4.4. These errors (ϵ_i) are often called residuals and represent the differences between observed (y) and estimated (y') variables. Ordinary least squares assumes identical weights (w_i) and independently distributed residuals with a normal distribution.

$$\min_{m,n} \sum_{i=1}^T w_i^2 \epsilon_i^2 \quad \text{where} \quad \epsilon_i = y_i - y'_i = yi - (mx_i + n) \quad (4.4)$$

It is frequent to observe that some residuals might have higher variance than others, meaning that those observations are effectively less certain. To contemplate such variability, weighted linear squares (WLS) regression (see Equation 4.4) applies a weighting function to the residuals. The confidence of the computed resistance index (observed variable) relies on the number of susceptibility test records manipulated. Hence, the sigmoid function has been used to define weights proportional to the population size.

Autoregressive Integrated Moving Average

An autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model which can be also applied in scenarios where data show evidence of non-stationarity. The autoregressive (AR) part expresses the variable of interest (resistance index) as a function of past values of the variable. The moving average (MA) indicates that the regression error is a linear combination of error terms which occurred contemporaneously and at various times in the past. An ARIMA(p,d,q) model is defined as shown in Equation 4.5, where p is the number of autoregressive terms, d is the number of differences needed for stationarity, q is the number of lagged forecast errors, and ϕ and θ are the coefficients of the model.

$$y'_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} - \sum_{j=1}^q \theta_j y_{t-j} \quad (4.5)$$

The interpretation of the parameter μ depends on the ARIMA model used for the fitting. In order to estimate the linear trend, it was interesting to consider exclusively MA models so that the expected value of μ was the mean of the one-time differenced series; that is, the slope coefficient of the un-differenced series. The Bayesian information criterion (BIC) was used to select the best ARIMA(0,1,q) model, being the one with the lowest BIC the preferred.

4.6.2 Analysis of the robustness of the methods

The process to generate a resistance time series signal from susceptibility data is defined by two parameters: shift and period. Regardless of the value of these parameters, the estimated trends should be independent of the granularity (shift) and show a consistent change when time spans overlap (period). In order to evaluate whether or not the regression methods provide coherent results for different shift_{period} configurations, the absolute difference between paired trends (SART distances) has been computed. The distribution of such distances is shown in Figure 4.8 for consecutive periods (left) and various granularities (right). Lower values indicate higher consistency in the estimation of trends.

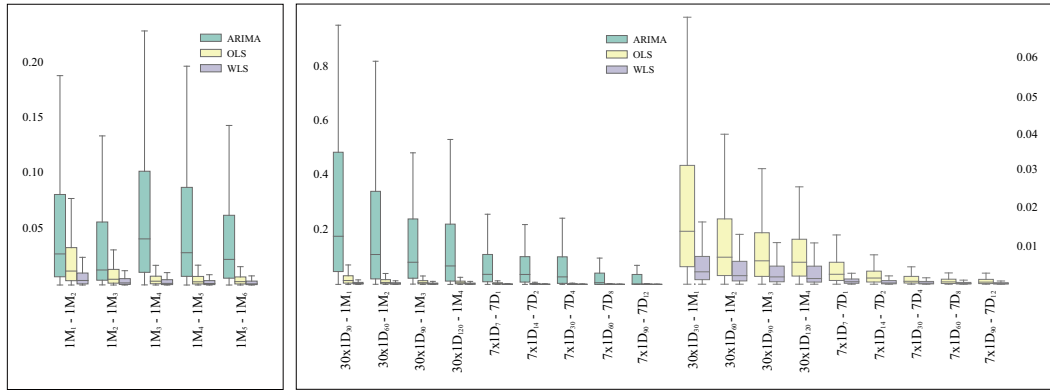


Figure 4.8: Distribution of paired Single Antimicrobial Resistance Trend (SART) distances. Comparison of ordinary least squares (OLS), weighted least squares (WLS) and autoregressive integrated moving average (ARIMA) in the following scenarios: consecutive periods (left) and equivalent granularities (right). A graph including exclusively OLS and WLS has been added on the latter to facilitate their comparison. The x-axis represents couplets of configurations to generate resistance time series under comparison.

Consistency of estimated trends for consecutive time spans

The length of the period determines the amount of susceptibility test records accounted to compute the resistance index. Lengthy periods provide smoother time series which are better approximated by the linear model, especially when overlapping time periods are considered. As a consequence, the SART distances decrease as shown by the median of the distributions (see left graph in Figure 4.8). This behaviour is consistent in OLS and WLS. However, it is worth highlighting the irregularities shown by ARIMA. The median and quartiles of the distributions indicate that WLS produces the most stable results and it is followed closely by OLS. Nonetheless, there is a considerable gap between these two methods and ARIMA. All the distances estimated by WLS were significantly smaller ($p < 0.001$) than those obtained using OLS and ARIMA.

Consistency of estimated trends for equivalent granularities

The SART measures ratio of change per time unit. Therefore, the monthly trend should be approximately four times the weekly trend and thirty times the daily trend. These correspondences are shown in Figure 4.8. Firstly, it is important to notice the substantial variation in the distribution of SART distances, which is one order of magnitude larger for ARIMA. Consequently, ARIMA has not been further considered for trend estimation. For the sake of clarity, the distribution of SART distances for OLS and WLS have been represented separately (see right graph in Figure 4.8). WLS presents the best performance in terms of granularity and the disparity with OLS is particularly visible for those scenarios in which independent time periods are used ($1M_1$ and $7D_1$). All the distances estimated by WLS were significantly smaller ($p < 0.001$) than those obtained using OLS and ARIMA.

To summarise, weighted least squares is robust against changes in the granularity of the generated time series. For this reason, the antimicrobial resistance trends obtained for three combinations is presented and compared with those reported by national and international organizations.

4.6.3 *Escherichia coli*: AMR in urine cultures

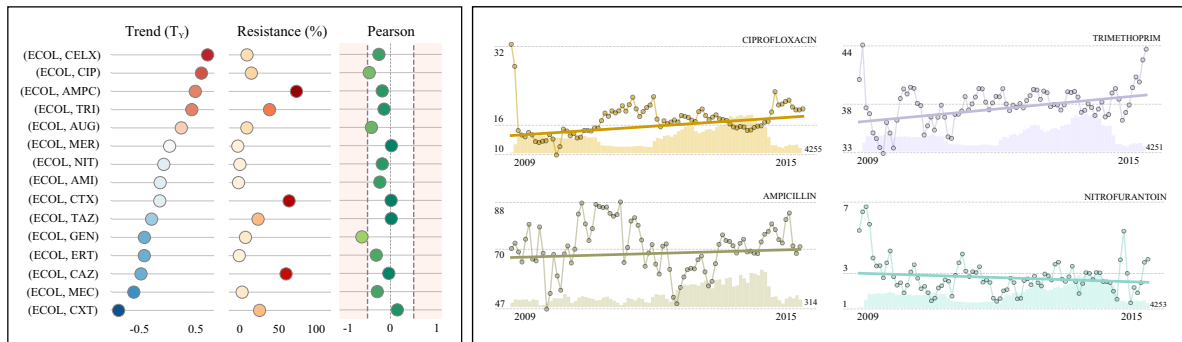
Escherichia coli (ECOL) is a Gram-negative bacteria and most strains are harmless being part of the normal flora of the gut. However, virulent strains can cause gastroenteritis, urinary tract infection or meningitis. *E. coli* is responsible for more than 85% of all urinary tract infections. There is an alarming resistance to ampicillin (69.6%), cefotaxime (60.8%), ceftazidime (57.3%) and trimethoprim (37.8%) with equivalent results in other studies (see Table 4.8). Furthermore, there has been a noticeable increase in resistance to cephalixin (0.7%), ciprofloxacin (0.6%) and trimethoprim (0.4%). Conversely, resistance to ampicillin (0.5%) and amoxicillin-clavulanate (0.2%) is positive yet not significant since the confidence intervals contain 0. Note the large variations in resistance to ampicillin along the years (see time series in Figure 4.8) which may be caused by external factors such as sample population at that period, advances in technology or policy changes. Nitrofurantoin is often identified as one of the most active agents to treat *E. coli* with resistance rates within the range 3.7-6% in 2003-2008 [37, 38], further stabilized to 3% in 2015-2017 [12, 14]. These rates harmonize with those presented in the corresponding resistance time series and the estimated marginally decreasing trend (-0.1%). While the resistance rate to nitrofurantoin (2.7%) is low, there are antimicrobials with even lower rates such as ertapenem (2.0%) or amikacin (1.1%). Furthermore, second-generation carbapenems show negligible resistance rates; meropenem (0.2%) and imipenem (0.2%).

Table 4.8: AMR summary for *Escherichia coli* in urine samples.

Antimicrobial	R(%) (95% CI)	References	T _M (%) (95% CI)	References	T _Y (%)	Pearson	Isolates
Cephalixin (CELX)	11.1 (10.9, 11.3)		0.055 (0.045, 0.065)		0.7 ↑	-0.25	79090
Ciprofloxacin (CIP)	16.3 (16.0, 16.5) [37, 39]		0.046 (0.031, 0.062) [14, 39]		0.6 ↑	-0.46	79239
Ampicillin (AMPC)	69.6 (68.2, 70.9) [39]		0.038 (-0.058, 0.134)		0.5 ↔	-0.18	4729
Trimethoprim (TRI)	37.8 (37.4, 38.1) [12][37][38][39]		0.033 (0.020, 0.046) [39]		0.4 ↑	-0.14	79133
Amoxicillin-Clavulanate (AUG)	10.9 (10.7, 11.2)		0.018 (-0.022, 0.059)		0.2 ↔	-0.42	79093
Meropenem (MER)	0.2 (0.1, 0.3)		0.002 (-0.002, 0.006)		0.0 ↔	0.02	9875
Nitrofurantoin (NIT)	2.7 (2.6, 2.8) [12][37][38][39]		-0.006 (-0.013, 0.001)		-0.1 ↔	-0.18	79108
Amikacin (AMI)	1.1 (0.9, 1.2)		-0.011 (-0.022, 0.000)		-0.1 ↔	-0.23	9786
Cefotaxime (CTX)	60.8 (59.9, 61.8)		-0.012 (-0.083, 0.059)		-0.1 ↔	0.01	9803
Tazocin (TAZ)	24.2 (23.3, 25.0) [37]		-0.023 (-0.078, 0.032)		-0.3 ↔	0.01	9878
Gentamicin (GEN)	9.3 (9.1, 9.5) [38]		-0.033 (-0.061, -0.005)		-0.4 ↓	-0.62	63399
Ertapenem (ERT)	2.0 (1.7, 2.3)		-0.033 (-0.050, -0.017)		-0.4 ↓	-0.31	8882
Ceftazidime (CAZ)	57.3 (53.3, 58.2)		-0.038 (-0.113, 0.037)		-0.5 ↔	-0.04	9810
Mecillinam (MEC)	5.4 (4.9, 5.8)		-0.048 (-0.071, -0.024)		-0.6 ↓	-0.29	9083
Cefoxitin (CXT)	26.0 (25.1, 26.8)		-0.069 (-0.123, -0.016)		-0.8 ↓	0.15	9798

Keys: CI=confidence interval; R=resistance; T_M=monthly trend; T_Y=yearly trend; ↑=significant increase; ↓=significant decrease.

Significance: A trend is significant if the CI does not include 0.

**Figure 4.9:** AMR summary for *Escherichia coli* in urine samples.

4.6.4 *Escherichia coli*: AMR in blood cultures

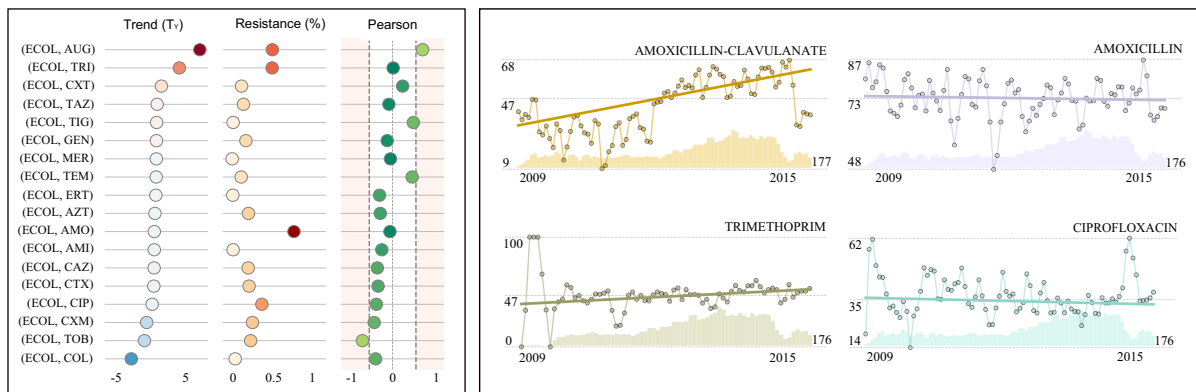
The national mandatory surveillance program has reported a consistent rise in the number of incidences of *E. coli* bacteremia in England [14]. Furthermore, the majority of antimicrobials under surveillance also presented an increase in resistance rates over the last years (see Table 4.9). Such rise is particularly significant for amoxicillin-clavulanate (4.3%) and trimethoprim (2.3%). The resistance trend for amoxicillin-clavulanate is significant yet the pearson coefficient indicates a high correlation between the number of records and the estimated resistance index. The introduction of MALDI-TOF mass spectrometry in 2011 might have caused this effect (see amoxicillin-clavulanate time series in Table 4.9). The high proportion of resistant isolates presented for these two antimicrobials (47.5% and 47.2% respectively) is only surpassed by amoxicillin (72.7%). Thus, there should be concerns on their use in clinical practice. Ciprofloxacin presents the fourth highest proportion of resistant isolates (35.2%) which has slightly decreased in recent years [26]. In contrast, surveillance in carbapenems shows negligible resistance rates which have remained constant over the years [14]. For instance, meropenem and ertapenem resistance rates (0.5% and 1.2%) and trends (0.0% and -0.1%) are shown below.

Table 4.9: AMR summary for *Escherichia coli* in blood samples.

Antimicrobial	R(%) (95% CI)	References	T _M (%) (95% CI)	References	T _Y (%)	Pearson	Isolates
Amoxicillin-clavulanate (AUG)	47.5 (45.8-49.2)	[12, 14]	0.359 (0.249, 0.470)	[12]	4.3 ↑	0.64	3317
Trimethoprim (TRI)	47.2 (45.4-49.1)		0.190 (0.079, 0.301)		2.3 ↑	0.01	2774
Cefoxitin (CXT)	11.5 (10.4-12.6)		0.041 (-0.006, 0.089)		0.5 ↔	0.22	3316
Tazocin (TAZ)	13.7 (12.6-14.9)	[12, 14, 40]	0.006 (-0.040, 0.052)	[12, 14, 40]	0.1 ↔	-0.08	3321
Tigecycline (TIG)	1.7 (1.2-2.2)		0.002 (-0.026, 0.030)		0.0 ↔	0.45	2734
Gentamicin (GEN)	16.6 (15.3-17.8)	[11, 14]	0.000 (-0.044, 0.045)	[11, 14]	0.0 ↔	-0.12	3322
Meropenem (MER)	0.5 (0.3-0.8)	[11, 12, 14, 40]	-0.001 (-0.020, 0.018)	[11, 12, 14, 40]	0.0 ↔	-0.05	3280
Temocillin (TEM)	11.1 (10.0-12.2)		-0.002 (-0.086, 0.082)		0.0 ↔	0.42	3044
Ertapenem (ERT)	1.2 (0.8-1.6)		-0.005 (-0.025, 0.016)		-0.1 ↔	-0.28	2992
Aztreonam (AZT)	19.6 (18.1-21.0)		-0.012 (-0.077, 0.052)		-0.1 ↔	-0.26	2925
Amoxicillin (AMO)	72.7 (71.2-74.2)		-0.017 (-0.085, 0.051)		-0.2 ↔	-0.06	3319
Amikacin (AMI)	1.6 (1.2-2.1)		-0.018 (-0.041, 0.006)		-0.2 ↔	-0.23	3044
Ceftazidime (CAZ)	19.3 (17.9-20.6)	[11, 26]	-0.019 (-0.065, 0.027)	[26]	-0.2 ↔	-0.33	3323
Cefotaxime (CTX)	20.3 (18.9-21.7)	[11, 26]	-0.021 (-0.070, 0.027)	[26]	-0.3 ↔	-0.31	3201
Ciprofloxacin (CIP)	35.2 (33.6-36.8)	[11]	-0.035 (-0.017, 0.037)		-0.4 ↔	-0.35	3320
Cefuroxime (CXM)	24.2 (22.8-25.7)		-0.080 (-0.137, -0.024)		-1.0 ↓	-0.39	3320
Tobramycin (TOB)	22.1 (20.6-23.6)		-0.099 (-0.188, -0.010)		-1.2 ↓	-0.65	2832
Colistin (COL)	4.0 (3.3-4.8)		-0.208 (-0.274, -0.141)		-2.5 ↓	-0.37	2606

Keys: CI=confidence interval; R=resistance; T_M=monthly trend; T_Y=yearly trend; ↑=significant increase; ↓=significant decrease.

Significance: A trend is significant if the CI does not include 0.

**Figure 4.10:** AMR summary for *Escherichia coli* in blood samples.

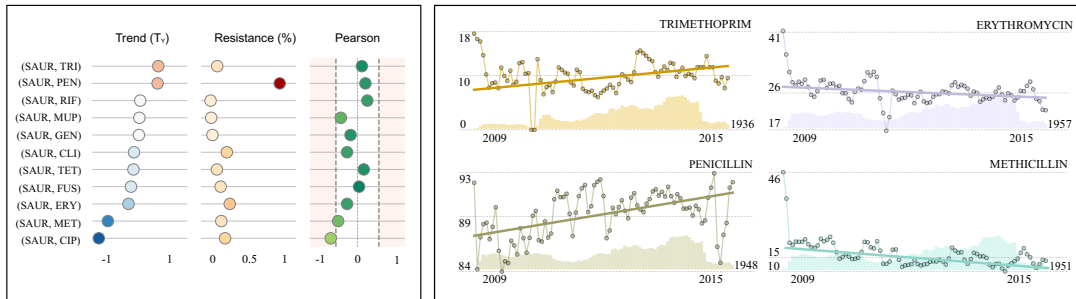
4.6.5 *Staphylococcus aureus*: AMR in wound cultures

Staphylococcus aureus (SAUR) is a Gram-positive bacteria typically found in the respiratory tract and the skin. It is a leading cause of bloodstream infections [41, 42], generally associated with breakages in the skin due to surgery, injury or use of intra-vascular devices such as catheters. Therefore, it is frequently acquired in hospitals [43]. Penicillin-resistant isolates were recognised in 1942 [44] reaching a proportion of 80% by late 1960s. Nowadays, the resistance rate to penicillin (89.4%) is the highest and has shown an increasing trend (0.6%) over the last years. There was also an emergence of methicillin-resistant *Staphylococcus aureus* (MRSA). It was first reported in a British hospital and became a worldwide problem in clinical medicine [26] with a peak of 43% in 2001 [45]. The Department of Health in England made reduction in rates of MRSA a priority with improvement of surveillance as one of their first actions. This led to a decrease in the number of resistant cases reported [46]. This decrease continued in 2012-2015 [12, 47] and coincides with the negative trend (-1.1%) presented in Table 4.10. Nowadays, approximately 15.3% of isolates are methicillin-resistant [12]. Also, there should be concerns on antimicrobials such as erythromycin (26.0%) and clindamycin (22.4%) with higher resistance rates and no further evidence of improvement. Moreover, while resistance rates to trimethoprim are not very high (10.1%), it has presented a noticeable rise (0.6%) in the last years. Since trimethoprim is clinically valuable to treat skin and soft tissue infections caused by MRSA, such rise constitutes a potential threat. On the other hand, rifampicin-containing treatments are known to improve outcomes in Staphylococcal wound infections presenting the lowest resistance rate (1.7%) and a constant trend (0.0%) [48].

Table 4.10: AMR summary for *Staphylococcus aureus* in wound samples.

Antimicrobial	R(%) (95% CI)	References	T _M (%) (95% CI)	References	T _Y (%)	Pearson	Isolates
Trimethoprim (TRI)	10.1 (9.8, 10.4)		0.052 (0.026, 0.077)		0.6 ↑	0.10	33525
Penicillin (PEN)	89.4 (89.1, 89.7)		0.050 (0.034, 0.065)		0.6 ↑	0.19	39901
Rifampicin (RIF)	1.7 (1.5, 1.8) [48]		0.001 (-0.015, 0.017) [48]		0.0 ↔	0.23	35141
Mupirocin (MUP)	2.5 (2.3, 2.6)		-0.001 (-0.020, 0.017)		0.0 ↔	-0.39	33716
Gentamicin (GEN)	4.1 (3.9, 4.3)		-0.003 (-0.023, 0.018)		0.0 ↔	-0.16	35255
Clindamycin (CLI)	22.4 (22.0, 22.8)		-0.016 (-0.034, 0.001)		-0.2 ↔	-0.24	39962
Tetracycline (TET)	9.7 (9.4, 10.0)		-0.018 (-0.041, 0.004)		-0.2 ↔	0.15	35429
Fusidic acid (FUS)	14.5 (14.2, 14.9)		-0.025 (-0.044, -0.006)		-0.3 ↓	0.04	39918
Erythromycin (ERY)	26.0 (25.6, 26.5)		-0.032 (-0.049, -0.015)		-0.4 ↓	-0.24	39971
Methicillin (MET)	15.3 (14.9, 15.7) [12]		-0.090 (-0.113, -0.068) [12]		-1.1 ↓	-0.45	39950
Ciprofloxacin (CIP)	20.1 (19.7, 20.5)		-0.116 (-0.156, -0.075)		-1.4 ↓	-0.62	35227

Keys: CI=confidence interval; R=resistance; T_M=monthly trend; T_Y=yearly trend; ↑=significant increase; ↓=significant decrease.
Significance: A trend is significant if the CI does not include 0.


Figure 4.11: AMR summary for *Staphylococcus aureus* in wound samples.

4.7 DISCUSSION

Antimicrobial surveillance is the cornerstone to strengthen the knowledge and evidence base to assess the burden of antimicrobial resistance. For this reason, international health care organizations have instigated research to understand the causes and consequences of AMR [19]. The conducted research, which gathered valuable information on resistance rates, incidences, prevalence and resistance trends, reported disparate levels of resistance around the world [1] and also within the UK [11–14]. However, there are significant gaps in the information available. Thus, it is necessary to develop stronger strategies to streamline antimicrobial surveillance and networks to facilitate information sharing.

4.7.1 *Susceptibility testing: behaviour and guidelines*

The number of susceptibility tests recorded among combinations varies considerably. The disparity among pathogens is induced by the hospital occurrence rate; leading to higher number of tests for pathogens which are a common cause of infection in hospital patients. On the other hand, hospital guidelines promote susceptibility testing on a small group of antimicrobials which are likely to be effective, causing the corresponding disparity. Surveillance outcomes are valuable to guide and support antimicrobial therapy selection for combinations that have been thoroughly studied. Unfortunately, the number of susceptibility tests available for a large proportion of combinations is insufficient to draw reliable conclusions. Hence, identifying changes in the effectiveness of antimicrobials could serve as biased evidence to motivate the inclusion of potentially hazardous combinations into existing guidelines.

4.7.2 *Advantages of overlapping time intervals in surveillance*

Antimicrobial surveillance is performed at different levels (e.g. local or national) and it is greatly affected by the size of the dataset considered. Versatile yet efficient analytic methods are required in those scenarios where data access or availability is restricted, such as clinical research. The main advantages of overlapping time intervals are: (i) it is a flexible approach which enables to adjust granularity and accuracy (ii) resulting time series are visually more legible and insightful (iii) enables the study of short-time variations in contrast to current mechanisms using sparse data points (years apart) and (iv) the outcomes are more consistent. On the other hand, it might originate certain relationship between consecutive observations as data is partially shared. Overall, this step is optional and might be particularly useful in scenarios where data is limited but decent levels of granularity are still required.

4.7.3 *Regression analysis: benefits and drawbacks*

Ordinary least squares (OLS) is perhaps the most popular method for trend estimation and has shown consistent results in our study. However, it is known to be greatly affected by outliers. To palliate this effect, weighted least squares (WLS) has been considered to reduce the contribution of outliers by considering the number of susceptibility tests available. Autoregressive integrated moving average (ARIMA) is a very popular suite of models which has proven to be robust in short-term forecasting. However, it has

two main limitations: (i) requires stationary time series and (ii) parameter tuning is not straightforward. Altogether, WLS was selected as the preferred method for trend estimation since it presented the best performance. Furthermore, it is easy to understand and implement increasing its practicability.

4.7.4 *The importance of surveillance data*

Despite global antimicrobial surveillance becoming a priority in recent years, homogeneity of antimicrobial policies does still produce different antimicrobial resistance outcomes [49]. For instance, it is widely documented that resistance rates are considerably higher in London than in the rest of UK, emphasizing the significance of local AMR surveillance. Health care organizations benefit from data on rates of antimicrobial resistance in many different ways: (i) contributes to the evidence base used for formulation of treatment guidelines from national to hospital unit levels (ii) can be used to assess the effectiveness and impact of interventions and (iii) has a key role in detecting the emergence and spread of previously uncommon or completely novel types of resistance. Furthermore, AMR surveillance plays a major role in patient management by providing data that influences clinical decision-making. Since it guides antimicrobial selection for empirical treatment could be an important clinical aid at point of care. For such reason, this information will be integrated in Enhanced Personalized and Integrated Care for Infection Management at Point of Care (EPiC IMPOC), a modular intelligent decision support system which aims to assist clinicians at the different stages of the infection management pathway [50, 51].

4.7.5 *Limitations*

The data considered in this study was not collected purposely and therefore a number of issues might arise. Firstly, the number of susceptibility tests recorded among combinations varies considerably. In fact, the majority of combinations did not have enough observations to provide reliable results. For this reason, exclusively results obtained for combinations with more than one thousand observations were presented. In addition, no information was provided as to whether specimens were collected in the hospital or primary care. Note that a mix of hospital-acquired and community acquired cases might difficult the interpretation of the results. Moreover, no additional information such as the underlying disease or the patient outcome was provided to support the susceptibility test data. This information is extremely useful to curate the data appropriately. For instance, microbiology cultures might contain pathogens which are not causing and infection and therefore present a conflicting AMR patterns.

Regarding pathogen and antimicrobial identification, a code system was included in the data. The majority of pathogens were identified at the species level. However, sometimes this information was provided at genus level. This is a clear limitation since different species from a common genus often present different resistant patterns. As such, the results provided at genus level could be biased. Similarly, external factors such as changes in susceptibility testing policies (e.g. MIC breakpoints), technology or outbreaks. For instance, the introduction of MALDI-TOF mass spectrometry in 2011 might have affected the number of microbiology tests requested by clinicians. However, while testing relies on hospital policies, suspicion of infection was assumed for all microbiology tests requested.

The provision of additional information such as the underlying disease, related vital signs and symptoms or the patient outcome could be useful to curate the data to obtain more specific and therefore informative results. However, the focus of this chapter is not to provide information about an specific condition but rather to delineate a generic methodology that can be applied to any susceptibility test dataset to extract AMR related information.

4.8 CONCLUSIONS

Surveillance is the cornerstone for assessing the burden of antimicrobial resistance and strengthens knowledge for action in support of stewardship program strategies by improving existing guidelines. The efficient use of susceptibility data provided by the overlapping time spans drops the dependence between granularity and accuracy of traditional surveillance systems. The robustness of weighted least squares regression facilitates resistance trend estimation and could be used to enhance existing surveillance systems which exclusively focus on resistant rates. Furthermore, there is an opportunity to investigate seasonal or other cyclic variations. Automating and facilitating access to antimicrobial surveillance reports through clinical decision support systems would enhance awareness among clinicians and possibly have an impact on antimicrobial prescription practices.

BIBLIOGRAPHY

- [1] Jim O'Neill. "Antimicrobial resistance: tackling a crisis for the health and wealth of nations". *Review on antimicrobial resistance* (2014), pp. 1–16.
- [2] Alison H Holmes, Luke SP Moore, Arnfinn Sundsfjord, Martin Steinbakk, Sadie Regmi, Abhilasha Karkey, et al. "Understanding the mechanisms and drivers of antimicrobial resistance". *The Lancet* 387.10014 (2016), pp. 176–187.
- [3] Jasper Littmann and AM Viens. "The ethical significance of antimicrobial resistance". *Public health ethics* 8.3 (2015), pp. 209–224.
- [4] C Pulcini, F Williams, N Molinari, P Davey, and D Nathwani. "Junior doctors' knowledge and perceptions of antibiotic resistance and prescribing: a survey in France and Scotland". *Clinical microbiology and infection* 17.1 (2011), pp. 80–87.
- [5] António Teixeira Rodrigues, Fátima Roque, Amílcar Falcao, Adolfo Figueiras, and Maria Teresa Herdeiro. "Understanding physician antibiotic prescribing behaviour: a systematic review of qualitative studies". *International journal of antimicrobial agents* 41.3 (2013), pp. 203–212.
- [6] Michael P Doyle, Guy H Loneragan, H Morgan Scott, and Randall S Singer. "Antimicrobial resistance: challenges and perspectives". *Comprehensive Reviews in Food Science and Food Safety* 12.2 (2013), pp. 234–248.
- [7] Esmita Charani, E Castro-Sanchez, N Sevdalis, Y Kyratsis, L Drumright, N Shah, et al. "Understanding the determinants of antimicrobial prescribing within hospitals: the role of "prescribing etiquette"". *Clinical Infectious Diseases* 57.2 (2013), pp. 188–196.
- [8] Esmita Charani, Rachel Edwards, Nick Sevdalis, Banos Alexandrou, Eleanor Sibley, David Mullett, et al. "Behavior change strategies to influence antimicrobial prescribing in acute care: a systematic review". *Clinical Infectious Diseases* 53.7 (2011), pp. 651–662.
- [9] Peter Davey, Erwin Brown, Esmita Charani, Lynda Fenelon, Ian M Gould, Alison Holmes, et al. "Interventions to improve antibiotic prescribing practices for hospital ins". *Cochrane Database Syst Rev* 4.4 (2013).
- [10] RJ Pinder, D Berry, A Sallis, and T Chadborn. *Antibiotic prescribing and behaviour change in healthcare settings: literature review and behavioural analysis*. Department of Health & Public Health England, 2015.
- [11] Public Health England. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR). Annual report 2014*. London: PHE; 2014.

-
- [12] Public Health England. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR). Annual report 2015*. London: PHE; 2015.
 - [13] Public Health England. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR). Annual report 2016*. London: PHE; 2016.
 - [14] Public Health England. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR). Annual report 2017*. London: PHE; 2017.
 - [15] European Centre for Disease Prevention and Control. *Surveillance of antimicrobial resistance in Europe 2013. Annual report of the European Antimicrobial Resistance Surveillance Network (EARS-Net)*. Stockholm: ECDC; 2014.
 - [16] European Centre for Disease Prevention and Control. *Surveillance of antimicrobial resistance in Europe 2014. Annual report of the European Antimicrobial Resistance Surveillance Network (EARS-Net)*. Stockholm: ECDC; 2015.
 - [17] European Centre for Disease Prevention and Control. *Surveillance of antimicrobial resistance in Europe 2015. Annual report of the European Antimicrobial Resistance Surveillance Network (EARS-Net)*. Stockholm: ECDC; 2016.
 - [18] European Centre for Disease Prevention and Control. *Surveillance of antimicrobial resistance in Europe 2016. Annual report of the European Antimicrobial Resistance Surveillance Network (EARS-Net)*. Stockholm: ECDC; 2017.
 - [19] World Health Organization. *Global antimicrobial resistance surveillance system (GLASS) report: early implementation 2016-2017*. Geneva: WHO; 2017. Licence: CC BY-NC-SA 3.0 IGO.
 - [20] Sunil Nair, Douglas Hsu, and Leo Anthony Celi. "Challenges and Opportunities in Secondary Analyses of Electronic Health Record Data". *Secondary Analysis of Electronic Health Records*. Cham: Springer International Publishing, 2016, pp. 17–26.
 - [21] Department of Health. *UK five year antimicrobial resistance strategy: 2013 to 2018*. London; 2013.
 - [22] Donald A Goldmann, Robert A Weinstein, Richard P Wenzel, Ofelia C Tablan, Richard J Duma, Robert P Gaynes, et al. "Strategies to prevent and control the emergence and spread of antimicrobial-resistant microorganisms in hospitals: a challenge to hospital leadership". *The Journal of the American Medical Association (Jama)* 275.3 (1996), pp. 234–240.
 - [23] Joel C Boggan, Ann Marie Navar-Boggan, and Ravi Jhaveri. "Pediatric-specific antimicrobial susceptibility data and empiric antibiotic selection". *Pediatrics* 130.3 (2012), e615–e622.
 - [24] Krumperman, PH. "Multiple antibiotic resistance indexing of *Escherichia coli* to identify high-risk sources of fecal contamination of foods." *Applied and Environmental Microbiology* 46.1 (1983), pp. 165–170.
 - [25] Ramanan Laxminarayan and Keith P Klugman. "Communicating trends in resistance using a drug resistance index". *BMJ open* 1.2 (2011), e000135.
 - [26] Alan P Johnson. "Surveillance of antibiotic resistance". *Philosophical Transactions of the Royal Society: Biological Sciences* 370.1670 (2015), p. 2014.0080.

- [27] GM Rossolini and E Mantengoli. "Antimicrobial resistance in Europe and its potential impact on empirical therapy". *Clinical Microbiology and Infection* 14.s6 (2008), pp. 2–8.
- [28] Public Health England. *UK standards for microbiology investigations: quality and consistency in clinical laboratories*. London: PHE; 2014.
- [29] British Society for Antimicrobial Chemotherapy. *BSAC methods for antimicrobial susceptibility testing. Version 12*. London: BSAC; 2013.
- [30] Jacques Acar. "Broad-and narrow-spectrum antibiotics: An unhelpful categorization". *Clinical Microbiology and Infection* 3.4 (1997), pp. 395–396.
- [31] Jeffrey S Gerber, Adam L Hersh, Matthew P Kronman, Jason G Newland, Rachael K Ross, and Talene A Metjian. "Development and application of an antibiotic spectrum index for benchmarking antibiotic selection patterns across hospitals". *infection control & hospital epidemiology* 38.8 (2017), pp. 993–997.
- [32] Paul R Rhomberg, Ronald N Jones, et al. "Antimicrobial spectrum of activity for meropenem and nine broad spectrum antimicrobials: report from the MYSTIC Program (2002) in North America". *Diagnostic microbiology and infectious disease* 47.1 (2003), pp. 365–372.
- [33] Moira Joelle Talpaert, Guduru Gopal Rao, Ben Symons Cooper, and Paul Wade. "Impact of guidelines and enhanced antibiotic stewardship on reducing broad spectrum antibiotic usage and its effect on incidence of Clostridium difficile infection". *Journal of Antimicrobial Chemotherapy* 66.9 (2011), pp. 2168–2174.
- [34] Joan L Slonczewski and John W Foster. *Microbiology: An Evolving Science: Third International Student Edition*. WW Norton & Company, 2013. Chap. 27.
- [35] Michael S Niederman. "Use of broad-spectrum antimicrobials for the treatment of pneumonia in seriously ill patients: maximizing clinical outcomes and minimizing selection of resistant organisms". *Clinical infectious diseases* 42.Supplement 2 (2006), S72–S81.
- [36] Rick van Saene, Sandy Fairclough, and Andy Petros. "Broad-and narrow-spectrum antibiotics: a different approach". *Clinical Microbiology and Infection* 4.1 (1998), pp. 56–57.
- [37] DJ Farrell, I Morrissey, D De Rubeis, M Robbins, and DAUK Felmingham. "A UK multicentre study of the antimicrobial susceptibility of bacterial pathogens causing urinary tract infection". *Journal of Infection* 46.2 (2003), pp. 94–100.
- [38] David C Bean, Daniel Krahe, and David W Wareham. "Antimicrobial resistance in community and nosocomial Escherichia coli urinary tract isolates, London 2005–2006". *Annals of clinical microbiology and antimicrobials* 7.1 (2008), p. 13.
- [39] Gunnar Kahlmeter, Jenny Åhman, and Erika Matuschek. "Antimicrobial resistance of Escherichia coli causing uncomplicated urinary tract infections: a European update for 2014 and comparison with 2000 and 2008". *Infectious diseases and therapy* 4.4 (2015), pp. 417–423.
- [40] Public Health Wales. *Antibacterial resistance in Wales: 2006-2015*. Cardiff: PHW; 2015.

-
- [41] Ad C Fluit, Mark E Jones, Franz-Josef Schmitz, Jacques Acar, Renu Gupta, Jan Verhoef, et al. "Antimicrobial susceptibility and frequency of occurrence of clinical blood isolates in Europe from the SENTRY antimicrobial surveillance program, 1997 and 1998". *Clinical Infectious Diseases* 30.3 (2000), pp. 454–460.
- [42] Hilmar Wisplinghoff, Tammy Bischoff, Sandra M Tallent, Harald Seifert, Richard P Wenzel, and Michael B Edmond. "Nosocomial bloodstream infections in US hospitals: analysis of 24,179 cases from a prospective nationwide surveillance study". *Clinical infectious diseases* 39.3 (2004), pp. 309–317.
- [43] Stephanie J Dancer. "Importance of the environment in meticillin-resistant *Staphylococcus aureus* acquisition: the case for hospital cleaning". *The Lancet infectious diseases* 8.2 (2008), pp. 101–113.
- [44] Charles H Rammelkamp and Thelma Maxon. "Resistance of *Staphylococcus aureus* to the Action of Penicillin." *Experimental Biology and Medicine* 51.3 (1942), pp. 386–389.
- [45] Alan P Johnson, Andrew Pearson, and Georgia Duckworth. "Surveillance and epidemiology of MRSA bacteraemia in the UK". *Journal of Antimicrobial Chemotherapy* 56.3 (2005), pp. 455–462.
- [46] Alan P Johnson, John Davies, Rebecca Guy, Julia Abernethy, Elizabeth Sheridan, Andrew Pearson, et al. "Mandatory surveillance of methicillin-resistant *Staphylococcus aureus* (MRSA) bacteraemia in England: the first 10 years". *Journal of antimicrobial chemotherapy* (2012), dkr561.
- [47] S Gerver, M Sinnathamby, S Bou-Antoun, S Kauser, M Canvin, J Abernethy, et al. "Annual epidemiological commentary: mandatory MRSA, MSSA and *E. coli* bacteraemia and *C. difficile* infection". *London: Public Health England* (2014).
- [48] Department of Health & Public Health England. *Antimicrobial resistance empirical and statistical evidence-based*. London: PHE; 2016.
- [49] Luke S. P. Moore, Rachel Freeman, Mark J. Gilchrist, Myriam Gharbi, Eimear T. Brannigan, Hugo Donaldson, et al. "Homogeneity of antimicrobial policy, yet heterogeneity of antimicrobial resistance: antimicrobial non susceptibility among clinical isolates from primary, secondary and tertiary care patients in London". *Journal of Antimicrobial Chemotherapy* 69.12 (2014), p. 3409.
- [50] Bernard Hernandez, Pau Herrero, Timothy M. Rawson, Luke S. P. Moore, Esmira Charani, Alison H. Holmes, et al. "Data-driven Web-based Intelligent Decision Support System for Infection Management at Point-Of-Care: Case-Based Reasoning Benefits and Limitations". 5 (2017), pp. 119–127.
- [51] Bernard Hernandez, Pau Herrero, Timothy Miles Rawson, Luke S. P. Moore, Benjamin Evans, Christofer Toumazou, et al. "Supervised learning for infection risk inference using pathology data". *BMC Medical Informatics and Decision Making* 17.1 (Dec. 2017), p. 168.

CASE-BASED REASONING

This chapter describes research towards the implementation of a decision support module to assist physicians in antimicrobial therapy selection. For such purpose, the chapter provides a brief introduction to the case-based reasoning (CBR) methodology (section 5.1). After this, the methodology used to generate the CBR models is explained; from the definition of distance (subsection 5.2.1) and similarity metrics (subsection 5.2.3) to the evaluation of the performance of the models in the context of information retrieval (subsection 5.2.5) and classification (subsection 5.2.6) problems. In particular, three experiments have been undertaken: an exceedingly simple scenario using the iris flower dataset (section 5.3), an imbalance scenario using the *Escherichia coli* protein location dataset (section 5.4) and an antimicrobial therapy selection trial within the intensive care unit to evaluate the system (section 5.5). To conclude, findings are discussed (section 5.6) and the main conclusions summarised (section 5.7).

5.1 THE CASE-BASED REASONING METHODOLOGY

The aim of CBR is to solve new problems based on the solutions of similar past problems in form of cases. CBR is considered as a methodology to follow rather than an algorithm in itself. Initially, the CBR methodology requires a set of cases or training examples denoted as case-base. Then, for a specific query, the CBR cycle is divided in four different phases (see Figure 5.1). The first phase *retrieves* from the case-base those cases that are relevant based on a predefined utility measure. In the second phase, advice is provided by *reusing* solutions from retrieved cases. This phase might involve adapting or combining solutions to fit the target problem. Subsequently, the query case is monitored to assess the validity of the solution applied. Finally, the outcome is *revised* to decide whether or not to *retain* the case in the case-base considering its re-usability. This methodology has been applied successfully in a wide variety of domains such as computer-aided design [1], fault diagnosis [2], bankruptcy prediction [3] and decision support [4]. Moreover, they have been used in the medical domain to develop clinical decision support systems to support tasks such as diagnosis, tutoring or planning. Some examples are diagnosis of stress-related disorders [5] or support for antimicrobial therapy selection [6].

In clinical environments, physician reasoning is based on knowledge acquired from past cases personally experienced. For such reason, it has been selected to be incorporated in the decision support system and strongly influenced the design of the user interface (see section 3.3).

As mentioned previously, there are four main phases within the CBR methodology which is denoted as the CBR cycle (see Figure 5.1). Lets review and discuss these phases in more detail.

- *Retrieve*: given a query, this phase retrieves from the database those cases that are relevant. The most common measure of utility is the similarity (or distance) between the cases. A critical decision in CBR systems resides on the selection of an appropriate distance metric [8]. In addition to the utility metric, it is required to define the number of cases to be retrieved from the case base.

- *Reuse*: propose a suitable solution for the query case by adapting and/or combining the solution of the retrieved cases. This phase is probably the most challenging within medical applications since it is almost impossible to generate adaptation/combination rules that consider all the important differences that may exist within the problem domain [9]. In general, these rules are rather specific for the problem domain.

- *Revise*: in this phase experts monitor the evolution of the case to assess whether or not the adapted/proposed solution was appropriate. The aim of this phase is to detect faulty adaptation/combination rules and incorporate expert knowledge into the CBR system to provide better solutions.

- *Retain*: determine whether or not the case is of utility to be added into the case-base.

In medical fields, attempts to apply the complete CBR cycle are rather exceptional. As mentioned previously, *adaptation* is extremely challenging in clinical domains. Thus, this research focuses only on the retrieval of similar cases to present them as information to the physicians. The motivation for abandoning the adaptation task is two-fold: in health-related application domains it is too complicated or even impossible to acquire sufficient adaptation knowledge but also the physicians tend to be interested in getting information about former similar cases, but prefer to reason about the current situation and the appropriate antimicrobial therapy prescription themselves.

5.1.1 The use of prototypes

A prototype is a generalised case that represents a set of individual cases that are very similar to each other; that is, the centroid of a cluster of individual cases. They were considered as an interesting technique to structure the case-base and to fill the knowledge gap between single cases and general knowledge [6]. While they never became a hot topic within the CBR community, they are used rather regularly in medical applications because they correspond to the reasoning of doctors in a natural way.

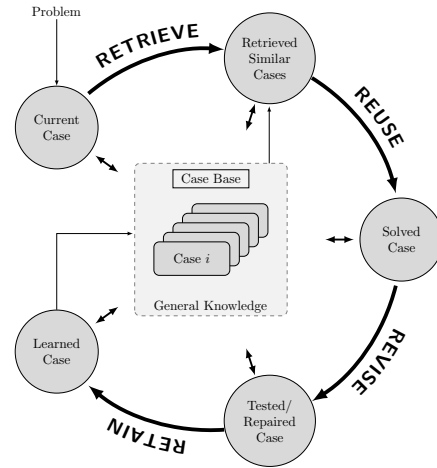


Figure 5.1: The CBR cycle. Diagram showing the different phases for a cycle within the case-based reasoning methodology as outlined by Aamodt and Plaza [7].

5.2 MATERIALS AND METHODS

The methodology followed in this chapter to create the CBR models is described graphically in Figure 5.2. In the first step, those observations which are duplicated or erroneous are discarded. After this, the remaining data is divided into complete observations (F_6) and incomplete observations ($\{F_n\}_{n=1}^5$). The complete observations are used to generate the CBR models. First, 25% of the observations in F_6 are isolated in the hold-out set (HOS) to further evaluate the translational utility. The remaining observations are used to generate and preliminarily evaluate the CBR models.

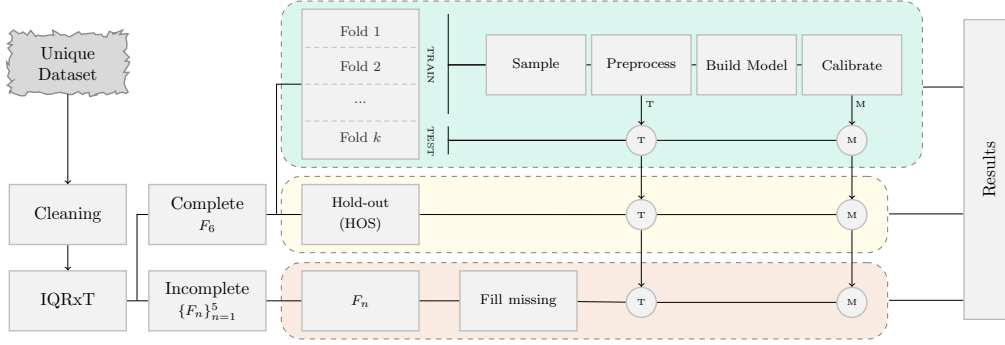


Figure 5.2: High-level methodology diagram for model creation and evaluation. First, data cleaning is performed. The remaining observations are grouped as complete or incomplete profiles. The former is further split into Cross-Validation Set (CVS) and Hold-out Set (HOS). Ten-Fold Stratified Cross-Validation is performed on CVS and two outputs are obtained in this step: a preprocessing equation to transform new observations (T) and a CBR model (M). It is important to highlight that preprocessing is performed using the train set. Finally, the performance of the CBR models is evaluated in HOS.

In this research, ten-fold stratified cross-validation has been used to assess how well the classifiers will generalize to an independent data set. This method divides the observations within the cross-validation set (CVS) in two categories: train and test. It is a common malpractice to perform data sampling on both categories. In this research, exclusively the train set is sampled, preprocessed and used to build the model (see Figure 5.2). From this procedure, a CBR model and a preprocessing equation (T) to format unseen observations are obtained. Finally, to assess the translational utility of these results the CBR models are validated using HOS. Note that these observations were not seen during data preprocessing and model training.

5.2.1 Definition of generic distance metrics

The most common measure of utility is the similarity (or distance) between the cases. These cases are often defined by a number of features which are arranged into a vector. Thus, the input parameters to the distance metrics (see Table 5.1) are two one-dimensional numeric vectors u and v . In addition, the parameter w_i indicates the weight associated to feature i . The presented distance metrics are applicable to numeric data yet further distances exist to compare other type of data such as boolean (e.g. jaccard similarity) or categorical (e.g. hamming distance) features.

Table 5.1: Distance metrics: description and equation.

Metric	Description	Equation
Manhattan	It computes the sum of the horizontal and vertical distances between points on a grid. It is also denoted city-block. Note that it is equivalent to the minkowski distance with $p=1$.	$d(u, v) = \sum_i w_i u_i - v_i $
Euclidean	It computes the "ordinary" straight-line distance between two points in the euclidean space. Note that is equivalent to the minkowski distance with $p=2$.	$d(u, v) = \left(\sum_i w_i u_i - v_i ^2 \right)^{1/2}$
Minkowski	It computes the distance between two points in a normed vector space. This metric is a generalization of euclidean and manhattan distances.	$d(u, v) = \left(\sum_i (w_i u_i - v_i ^p) \right)^{1/p}$
Canberra	It computes the distance between pairs of points in a vector space. This metric is a weighted version of the manhattan distance.	$d(u, v) = \sum_i w_i \frac{ u_i - v_i }{ u_i + v_i }$
Braycurtis	It computes the distance between vectors based on the counts at each site. The range of the metric is $[0,1]$ if all coordinates are positive.	$d(u, v) = \frac{\sum_i w_i u_i - v_i }{\sum_i w_i u_i + v_i }$
Correlation [†]	It computes the correlation between two vectors where $\bar{u}$ is the mean of the elements of u and $u \cdot v$ is the dot product of u and v .	$d(u, v) = 1 - \frac{(u - \bar{u}) \cdot (v - \bar{v})}{\ (u - \bar{u})\ _2 \ (v - \bar{v})\ _2}$
Cosine [†]	It computes the distance between vectors of an inner product space that measures the cosine of the angle between them where $u \cdot v$ is the dot product of u and v .	$d(u, v) = 1 - \frac{(u \cdot v)}{\ u\ _2 \ v\ _2}$
Chebysev	It computes the distance between two vectors as the greatest of the differences along any coordinate dimension. It is also denoted maximum metric or L_∞ .	$d(u, v) = \max_i w_i (u_i - v_i) $

[†] indicates that weights can be passed to compute the norm ($\|x\|_2$) and the dot product ($x \cdot y$).

5.2.2 Definition of case features: antimicrobial therapy selection in the ICU

The identification of parameters to define the CBR case was led by Luke Moore and Timothy Rawson. Firstly, an analysis of the variables which are key drivers of decisions around antimicrobial prescribing in augmented care was conducted by Luke Moore [10]. This analysis also evaluated the importance of these variables and how they are integrated into the decision making process. A critical analysis of the existing scientific literature on CDSS in the ICU and a qualitative analysis were undertaken. A total of 20 participants with different experience levels were recruited among both critical specialists (9/20) and infection specialists (11/20). Analysis of the semi-structured interview transcripts enabled description of the variables which directly impact infection-related decision making processes in critical care. A similar approach was conducted by Timothy Rawson to explore the clinical variables and the infection management pathway around antimicrobial prescribing in secondary care [11]. A total of 20 participants were recruited: on-rotation physicians (7/20), specialist trainees (4/20) and consultants (9/20). The study identified the pathway followed by clinicians to prescribe antimicrobials to follow a stepwise Bayesian-like approach [12]. The two infection specialists (Luke

Moore and Timothy Rawson) reviewed the identified variables and corroborated the findings. To conclude, relevant literature was reviewed to provide further evidence in support of these selected variables (see appendix E.1). Identified variables were compared to available data that was electronically available at the time of development to inform the design of the CBR decision support module.

To monitor the evolution of the patient, six variables were eventually selected based on their availability within the electronic health records and their use in infection management. These variables were C-reactive protein (CRP), white cell count (WCC), creatinine (CRE), alanine aminotransferase (ALT), bilirubin (BIL) and alkaline phosphatase (ALP) [13–18]. Lactate was also identified to be an important blood marker, however this biomarker was not routinely available for the majority of patients within the electronic database [19–21]. Physiological parameters such as heart rate, respiratory rate or blood pressure were not available electronically [22, 23].

5.2.3 Definition of similarity metrics: antimicrobial therapy selection in the ICU

The CBR methodology retrieves cases from the database based on a similarity measure. The CBR cases are defined by a compendium of variables (see appendix E.1) that can be divided in five different groups: metadata, description, solution, justification and result. The variables within the description group are then used to compute the similarity measure. Lets define a query case description as $X = \{x_1, \dots, x_i\}$ where x_i represents the attribute i . Since the type of attributes stored in a case vary (e.g. numbers, strings or sets) it is required to define partial similarity metrics that are applied at attribute-level (e.g. body temperature). The similarity value provided by such metrics ranges in the interval $[0,1]$ where 1 indicates maximum similarity. Later, to obtain the overall similarity between a query case (X) and any case from the database (Y) partial similarities for all attributes are combined using a weighted euclidean distance (see Equation 5.1).

$$S_{overall}(X, Y) = \sum_{i=1}^I w_i S_{partial}(x_i, y_i) \quad \text{where} \quad \sum_{i=1}^I w_i = 1 \quad (5.1)$$

The currently defined partial similarity metrics are explained below where x and y represent a feature from the CBR cases under comparison. Note that the application of certain similarity metrics is restricted to the type of the variable.

- *Equal similarity*: It returns maximum similarity when the inputed values are identical (see Equation 5.2) and it can be applied to any type of attribute.

$$S_{Equal}(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases} \quad (5.2)$$

- *Interval similarity*: Let's consider an interval $[a, b]$ with length L . This function measures the similarity between two numerical input values within the defined interval as shown in Equation 5.3.

$$S_{Interval}(x, y) = 1 - \frac{|x - y|}{L} \quad (5.3)$$

- *Table similarity*: This function requires a previously defined table containing the similarity score between pairs of values for a given attribute. Hence, it is applied to categorical attributes and looks up the similarity score in the predefined table (see Equation 5.4).

$$S_{Table}(x, y) = cell_{x,y} \quad (5.4)$$

Before introducing more complex functions to measure similarities, let's define two simple functions based on the previously defined table similarity S_{Table} . They receive two inputs: a single categorical query value (e_q) and a set of categorical values (E) respectively. The S_{Tmax} function returns the maximum table similarity found for the query value (e_q) and each of the values $e_i \in E$ (see Equation 5.5).

$$S_{Tmax}(e_q, E) = \max (S_{Table}(e, e_i)) \quad \forall e_i \in E \quad (5.5)$$

Similarly, the S_{Tavg} function returns the average similarity for the query value (e_q) and all the values in the enumerated set E (see Equation 5.6).

$$S_{Tavg}(e, E) = \frac{1}{|E|} \sum_{e_i \in E} S_{Table}(e, e_i) \quad (5.6)$$

- *Sets similarity*: A similarity table between different pairs of organs is provided by clinicians and microbiologists based on their experience. Let us assume we have a set of N possible different organs ($O = \{o_i\}_{i=0}^N$). The function receives two set of organs as inputs (O_1 and O_2) where $|O_1| \geq |O_2|$ and computes the similarity as shown in Equation 5.7.

$$S_{organ}(O_1, O_2) = \frac{\sum_{o \in \{O_1 \cap O_2\}} S_{Table}(o, o) + \sum_{o \in \{O_1 - O_2\}} S_{Tmax}(o, O_2)}{|O_2|} \quad (5.7)$$

- *Cluster Similarity*: Initially, infectious organisms are grouped and a similarity table for pairs of groups is provided by clinicians and microbiologists based on their experience. Let us assume we have a set of N possible different microorganisms ($M = \{m_i\}_{i=0}^N$) which are clustered in G different groups ($G = \{g_i\}_{i=0}^G$) using the grouping function $G(x_i) = g_z$. Hence, for two sets of microorganisms as input (M_1 and M_2) where $|M_1| \geq |M_2|$ the similarity is computed as shown in Equation 5.8.

$$S_{pathogen}(M_1, M_2) = \frac{S_{EM}(M_1, M_2) + S_{EG}(M_1, M_2) + S_{DG}(M_1, M_2)}{|EM| + |EG| + \min(1, |DG|)} \quad (5.8)$$

The first term counts the total number of Equal Microorganisms among the two sets; that is, the cardinality of the their intersection (see Equation 5.9).

$$S_{EM}(M_1, M_2) = |M_1 \cap M_2| = |EM|. \quad (5.9)$$

For the second term, identical microorganisms are subtracted from the original sets obtaining two new sets with exclusively Different Microorganisms (DM_i). Note this

two sets may contain microorganisms belonging to the same group. Hence, we also define two sets containing All Groups (AG_i) for the remaining microorganisms.

$$\begin{aligned} DM_1 &= M_1 \setminus M_1 \cap M_2 = M_1 - EM \\ DM_2 &= M_2 \setminus M_1 \cap M_2 = M_2 - EM \end{aligned} \quad (5.10)$$

$$\begin{aligned} AG_1 &= \{G(x_i)\} \quad \forall x_i \in DM_1 \\ AG_2 &= \{G(x_i)\} \quad \forall x_i \in DM_2 \end{aligned} \quad (5.11)$$

The second term counts the total number of Equal Groups once identical microorganisms have been discarded from the two original sets (see Equation 5.12). Since microorganisms in the same group are treated with similar antibiotics, if more than one microorganism is found for a specific group, it is considered only once.

$$S_{EG}(M_1, M_2) = |AG_1 \cap AG_2| = |EG| \quad (5.12)$$

The third term measures the average similarity for each group in the query input and all groups in the database input (see Equation 5.13). Note that all common groups are previously discarded as shown in Equation 5.14.

$$\begin{aligned} DG_1 &= AG_1 \setminus AG_1 \cap AG_2 = AG_1 - EG \\ DG_2 &= AG_2 \setminus AG_1 \cap AG_2 = AG_2 - EG \end{aligned} \quad (5.13)$$

$$S_{DG}(M_1, M_2) = \frac{\sum_{x_i \in DG_1} S_{Tavg}(x_i, DG_2)}{\max(|DG_1|, |DG_2|)} \quad (5.14)$$

To conclude, highlight that all the previously defined similarity metrics provide a value within the range [0, 1]. Thus, these similarity metrics can be easily converted to distance metrics as follows:

$$D(x, y) = 1 - S(x, y)$$

5.2.4 Assigning feature importance

In CBR, the weights of feature attributes directly affect the quality of problem solving. For such reason, the allocation method of the feature attribute weights is a very important research direction in the CBR model [24]. In recent years, the methods for weight allocation in CBR systems have attracted increasing attention and are based on previous knowledge [25], grid search [24], membrane computing [26] or genetic algorithms [27]. The different strategies used in this research to assign the weights are described in Table 5.2. The identification of weights was conducted by Luke Moore and Timothy Rawson based on the literature review and qualitative analyses previously described [10, 11]. The factors associated with decision making were ranked based on the mode from the order in which they were reported by individual physician. This allowed comparison and weighting of individual factors that were reported to influence the decision making process.

Table 5.2: Description of strategies to assign CBR feature weights.

Strategy	Description	Notes
Uniform	The features are weighted equally.	
Defined by infection specialists	Based on scientific evidence, a panel of expert infection specialists evaluated the clinical symptoms that help in the diagnosis and further treatment of infectious diseases. After this, they reached a consensus on the definition of the appropriate weight configuration according to scientific evidence but also their own expertise.	
Grid-search	This strategy trains and evaluates models with a predefined set of configurations using 10-stratified cross validation and selects the one that performed the best. The set configuration includes the combinations with repetition and permutations of the possible weight values in a vector v . Note that the resulting weight configurations are of length equal to the number of feature and must add up to one. The possible weight values (v) used were: [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]	$F_1 = 1$ $F_2 = 12$ $F_3 = 67$ $F_4 = 287$ $F_5 = 1002$ $F_6 = 3004$ $F_7 = 8009$

F_i indicates the total number of possible weight configurations for i features.

5.2.5 Performance metrics for information retrieval problems

The normalized Discounted Cumulative Gain (nDCG) score is often used in information retrieval to measure the quality of a ranking (e.g. search result). For that purpose, it measures the usefulness of a retrieved element based on the position in the produced ranking. Firstly, let's define the Discounted Cumulative Gain (DCG) score accumulated at a particular rank position p where rel_i is the graded relevance of the results at position i as shown in Equation 5.15.

$$DCG_p = \sum_{i=1}^p \frac{rel_i}{\log_2(i+1)} \quad (5.15)$$

This score penalizes highly relevant documents appearing lower in the search result list. However, the search result list might vary in length depending on the query. Thus, to compare the performance of queries with different lengths, the previous score needs to be normalized. This is done by sorting all relevant documents in the corpus by their relative relevance to produce the maximum possible DCG through position p , also called ideal DCG (see Equation 5.16).

$$IDCG_p = \sum_{i=1}^{|REL|} \frac{rel_i}{\log_2(i+1)} \quad (5.16)$$

Thus, the normalized score is computed as stated in Equation 5.17. The nDCG score is within the range [0,1] where 1 indicates perfect ranking. These values can be averaged for all queries to obtain a measure of the average performance ranking.

$$nDCG_p = \frac{DCG_p}{IDCG_p} \quad (5.17)$$

5.2.6 Performance metrics for classification problems

There are many different metrics for assessing the performance of classifiers [28, 29]. For binary classifiers, most of them are based on four simple measures: the number of true positives (TP), the number of false positives (FP), the number of true negatives (TN) and the number of false negatives (FN). Sensitivity, specificity and overall accuracy are commonly used to demonstrate classifiers performance [30–32]. Note however, that accuracy might not be appropriate when the class sizes differ considerably [33]. The Cohen’s Kappa statistic is considered a robust measure for evaluating the accuracy in multiclass problems [34]. Alternatively, a number of strategies to generalise all the binary performance metrics to multiclass domains can be used. The “micro” average computes the metric from the individual TP, FP, TN, FN of the categories. In the “macro” average, the performance of each individual category is averaged. The definition and equations of the previously mentioned metrics and the strategies to extend them to the problem of multiclass classification are shown in Table 5.3.

Table 5.3: Evaluation metrics: description and equation

Metric	Description	Equation
Sensitivity	Proportion of observed positives that are correctly identified as such (i.e. percentage of culture-positive profiles correctly identified as positive). Also called recall (REC) or true positive rate (TPR).	$SENS = \frac{TP}{TP+FN}$
Specificity	Proportion of observed negatives that are correctly identified as such (i.e. percentage of culture-negative profiles correctly identified as negative). Also called true negative rate (TNR).	$SPEC = \frac{TN}{TN+FP}$
“micro” average	Method to extend binary performance metrics to multiclass domains by computing the metric from the individual TP, FP, TN, FN of the categories. This strategy gives equal weight to each individual decision. Thus large classes will dominate small ones.	$SENS' = \frac{\sum_{n=1}^K TP_n}{\sum_{n=1}^K TP_n + \sum_{n=1}^K FN_n}$
“macro” average	Method to extend binary performance metrics to multiclass domains by averaging the value of the metric for each individual category. This strategy gives equal weight to every category.	$SENS' = \frac{SENS'_{n=1} + \dots + SENS'_{n=K}}{K}$

Keys: KAPPA=Cohen’s kappa coefficient; REC=recall; SENS=sensitivity; SPEC=specificity; TP=true positive; TN=true negative; FP=false positive; FN=false negative; TNR=true negative rate;

5.2.7 Overall performance of the CBR model

The aim of the CBR system implemented in this research is to return a ranked list of similar cases. So far, two configuration parameters of the CBR system have been described: the distance metric and the feature weights. However, an additional configuration parameter is required which refers to the number of elements returned in the ranked list (k). The maximum length of the ranked list (k_{max}) has been fixed to 20 and the previously defined evaluation metrics have been computed for every possible length of the ranked list ($k=1..k_{max}$). Finally, for a given metric, the average for all the lengths is computed to evaluate the overall performance. Note that the CBR system presents to physicians 20 cases yet it cannot ensure that all of them will be explored. For such reason, it is essential to rank relevant items first to promote that these results are seen. In addition, it is desirable to provide results which are consistent among different k .

5.2.8 Statistical analysis

Given the outputs of each iteration, the two-sample F-test was used to check that the samples come from independent distributions with equal variances. Subsequently, the corresponding t-test or corrected t-test (Welch's test) was applied to compare the behaviour of the models searching for statistically significant differences. The populations were assumed to have normal distributions. The final measure presented for each configuration of the CBR is the average of the 10-fold validation for all the metrics.

5.2.9 Outline of the experiments conducted

This section briefly outlines the experiments further described in this chapter. The first experiment (Experiment I) assesses the performance of CBR in an exceedingly simple scenario. For such purpose, the well-known iris flower dataset has been chosen. In addition, the second experiment (Experiment II) investigates the effect of imbalanced categories using the *Escherichia coli* protein location dataset. These two experiments assess the behaviour of the CBR methodology prior to the translation into clinical practice. To conclude, the last experiment (Experiment III) evaluates the performance of CBR for antimicrobial therapy selection. Since there was a lack of clinical data relating the patient medical record and the appropriate antimicrobial therapy, the CBR module was integrated into EPiC IMPOC and trialled directly in the intensive care unit.

5.3 EXPERIMENT I: A SIMPLE SCENARIO (IRIS FLOWER DATASET)

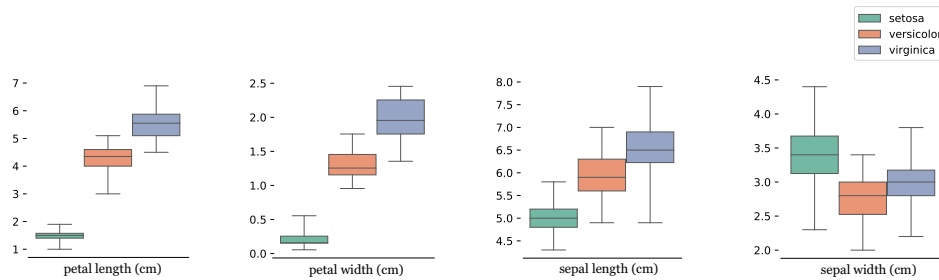
The aim of this experiment is to assess the performance of CBR to retrieve and rank appropriately observations from the database on an exceedingly simple scenario. For such purpose, one of the best known databases within the pattern recognition literature has been used. The iris flower dataset (see Table 5.4) has four features and contains three categories of 50 instances each, where each category refers to a type of iris plant. Thus, the categories are balanced. The dataset does not have missing values.

Table 5.4: Iris flower dataset description.

Features		Categories		Notes	
Petal length	continuous	Setosa	50 (33%)	Instance	150
Petal width	continuous	Versicolor	50 (33%)	Missing data	No
Sepal length	continuous	Virginica	50 (33%)	Imbalance categories	No
Sepal width	continuous				

5.3.1 Distributions of the iris flower dataset features

The density distribution of the features is presented in Figure 5.3 for each of the categories (*setosa*, *versicolor* and *virginica*) separately. The majority of the distributions are highly symmetric about the mean and can be roughly seen as normal distributions. One category (*setosa*) is linearly separable from the other two. Note that the mean of the distribution significantly differs from the other categories for petal length and petal width. Moreover, the standard deviation is also small. On the contrary, *setosa* and *versicolor* categories overlap on all the features. However, as occurred before, petal length and petal width seem to be the most informative features to discern among them. To conclude, note that the distributions for petal length and petal width are very similar among all categories. The correlation between these features is high indicating that most of the information is potentially redundant.

**Figure 5.3:** Distribution for iris categories. The distribution of the measurements for each of the features grouped in three categories: *setosa*, *versicolor* and *virginica*.

5.3.2 Analysis of the CBR retrieval performance

This section compares the performance of different CBR configurations to retrieve and appropriately rank by similarity observations from the database. Firstly, the effect of a number of distance metrics was explored under equal feature weights (see Tables 5.5 and 5.6). After this, a total of 286 different weights were evaluated using grid search (see Table 5.2). The top four CBR configurations were selected and included in Table 5.7. The

metrics included from left to right are: normalized discounted cumulative gain (nDCG) score, Cohen's kappa coefficient (KAPPA), sensitivity (SENS) and specificity (SPEC). For the sensitivity and specificity, micro/macro strategies to compute the scores on multi-class problems are presented. Since the CBR employs distance metrics, the features were normalized. The grey shade indicates for each scenario the model which provides the best performance.

Table 5.5: Comparison of distance metrics (I).

		nDCG	KAPPA	SENS	SPEC
UNI	MAN	0.918 [†]	0.89	0.93/0.93	0.96/0.96
	EUC	0.921 [†]	0.90	0.93/0.93	0.97/0.97
	MINK3	0.917 [†]	0.89	0.93/0.93	0.96/0.96
	MINK4	0.915 [†]	0.88	0.92/0.92	0.96/0.96
	MINK5	0.913 [†]	0.90	0.93/0.94	0.97/0.97

[†] indicates groups where difference is not significant.

Description: This table shows the performance of a CBR in a multiclass problem with three balanced categories. The weights are equal for all the features (uniform). The number of neighbors investigated was $k=1..20$.

Table 5.6: Comparison of distance metrics (II).

		nDCG	KAPPA	SENS	SPEC
UNI	BRAY	0.925	0.90	0.93/0.94	0.97/0.97
	CAN	0.906	0.88	0.92/0.92	0.96/0.96
	CHEB	0.908	0.90	0.93/0.94	0.97/0.97
	CORR	0.741 [‡]	0.66	0.77/0.78	0.89/0.88
	COS	0.760 [‡]	0.74	0.82/0.83	0.91/0.91

[‡] indicate groups where difference is not significant.

Description: This table shows the performance of a CBR in a multiclass problem with three balanced categories. The weights are equal for all the features (uniform). The number of neighbors investigated was $k=1..20$.

A critical decision in CBR systems resides on the selection of an appropriate distance metric [8]. For such reason, a number of well-defined distance metrics have been compared. The performances are quite homogeneous among the minkowski-based distances (see Table 5.5) and the negligible variations in the nDCG score are not statistically significant. Moreover, the only perceptible variation in performance is found for the correlation and the cosine metrics. Thus, although the best performing metric is the euclidean distance, the differences are not significant. Conversely, the configuration of the feature weights does affect the retrieval performance of CBR (see Table 5.7). For instance, the nDCG scores obtained for the top four weight configurations identified through grid search are significantly higher than those obtained using uniform weights. Similarly, the nDCG scores for $SEARCH_{i=1,2}$ are significantly better than those obtained for $SEARCH_{i=3,4}$. However, there is no statistical significance within groups. Thus, different weight configurations may present the same performance. Note that the uniform weights configuration was ranked 145th. To conclude, remark that $SEARCH_1$ discarded one of the redundant features (petal length) and still presents the best overall performance.

Table 5.7: Comparison of weight configurations.

		nDCG	KAPPA	SENS	SPEC	weights
EUC	UNI	0.921	0.90	0.93/0.93	0.97/0.97	[1/4, 1/4, 1/4, 1/4]
	SEARCH ₁	0.939 [‡]	0.92	0.95/0.95	0.98/0.97	[0.0, 0.4, 0.5, 0.1]
	SEARCH ₂	0.938 [‡]	0.90	0.93/0.93	0.97/0.97	[0.0, 0.5, 0.4, 0.1]
	SEARCH ₃	0.937 [†]	0.91	0.94/0.94	0.97/0.97	[0.1, 0.4, 0.4, 0.1]
	SEARCH ₄	0.936 [†]	0.91	0.94/0.94	0.97/0.97	[0.1, 0.5, 0.3, 0.1]

[‡] indicate groups where differences are not statistically significant.

Description: This table shows the performance of a CBR in a multiclass problem with three balanced categories. A total of 286 weights were evaluated in grid search; the top four are presented. The uniform weights (UNI) are shown as reference. The number of neighbors investigated was $k=1..20$.

The left graph in Figure 5.4 presents the evolution of the NDCG score through different lengths of the ranking list (k) for two different weight configurations (UNI and SEARCH₁). The performance is similar for small values of k . However, for larger values of k , SEARCH₁ is superior at presenting relevant observations ranked first in the retrieved list. The right graph in Figure 5.4 presents the evolution of the NDCG score for each category independently. The results obtained for the linearly separable category (*setosa*) are ideal since the maximum NDCG score is obtained. This means that the retrieved ranking for all the queries presented exclusively relevant observations (from the *Setosa* category). On the other hand, the performance of the other two non linearly separable categories is similar.

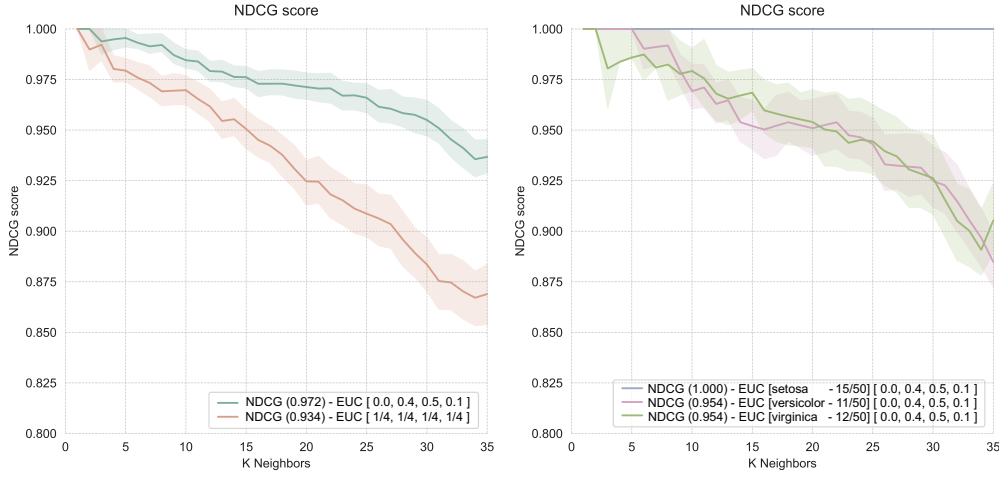


Figure 5.4: nDCG comparison On the left, evolution of the nDCG score through different k on two scenarios: (i) uniform weights (UNI) and (ii) best weight configuration found using grid search (SEARCH₁). On the right, the evolution of nDCG for each category independently. The information in the legend includes the name of the score (nDCG), the mean for all k , the distance metric, the number of query/-database cases and the feature weights.

It is important to highlight that the nDCG score evaluates the rank of the observations retrieved from the database while the rest of the metrics evaluate the final category assigned. Hence, the test observations were accurately classified for all the configurations (as indicated by the classification metrics) yet the rank of the observations retrieved from the database differed. However, in this exceedingly simple scenario, the rank did not greatly affect the final predicted category.

5.4 EXPERIMENT II: IMBALANCED CATEGORIES (ECOLI PROTEIN LOCATION DATASET)

The aim of this experiment is to assess the performance of CBR to retrieve and rank appropriately observations from the database on scenarios with imbalance categories. For such purpose, the *Escherichia coli* protein location database has been used. This dataset (see Table 5.8) has seven features; five of them continuous and two of them binary. The dataset contains eight categories with highly imbalanced classes. The dataset does not have missing values.

Table 5.8: *Escherichia coli* dataset description.

Features		Categories	Notes	
mcg	continuous	cytoplasm (cp)	143 (43%)	Instances
gvh	continuous	inner no-sequence (im)	77 (23%)	Missing data
lip	discrete	periplasm (pp)	52 (15%)	Imbalance categories
chg	discrete	Inner uncleavable-sequence (imU)	35 (10%)	Yes
aac	continuous	outer (om)	20 (6%)	
alm1	continuous	outer lipoprotein (omL)	5 (1%)	
alm2	continuous	inner lipoprotein (imL)	2 (1%)	
		Inner cleavable-sequence (imS)	2 (1%)	

Keys: *mcg*=McGeoch's method for signal sequence recognition; *gvh*=von Heijne's method for signal sequence recognition. *lip*=von Heijne's signal peptidase II consensus sequence score; *chg*=presence of charge on n-terminus of predicted lipoproteins; *acc*=score of discriminant analysis of the amino acid content of outer membrane and periplasmic proteins; *alm1*=score of the ALOM membrane spanning region prediction program; *alm2*=score of ALON program after excluding putative cleavable signal regions from the sequence; *cp*=cytoplasm; *im*=inner membrane without signal sequence; *pp*=periplasm; *imU*=inner membrane with uncleavable cytoplasm; *om*=outer membrane; *omL*=outer membrane lipoprotein; *imL*=inner membrane lipoprotein; *imS*=inner membrane with cleavable signal sequence;

5.4.1 Distributions of the *ecoli* protein location dataset features

The density distribution of the continuous features is presented in Figure 5.5 for each category. In particular, the values of *alm1* and *mcg* seem to be of utility to identify the *cp* class since the mean is clearly the lowest. Note that this category is the most common within the dataset. In these two features it is also visible that *imS* and *imL* present similar distributions regarding both median and standard deviation. However, the number of observations in these classes is not representative. In overall, there are no clear visible patterns among the features.

Note that the dataset contains two features that are discrete (*chg* and *lip*). In these cases, the proportion of observations has been presented. For instance, the *lip* feature has two possible values; either 0.48 or 1.0. Hence from the corresponding graph (see Figure 5.5) the proportion of observations with a value of 0.48 are as follows: 100% (*cp*), 98% (*im*), 0% (*im*), 100% (*imS*), 96% (*imU*), 94% (*om*) and 100% (*omL*). Note that this representation can also be used for discrete or categorical variables with more than two possible values.

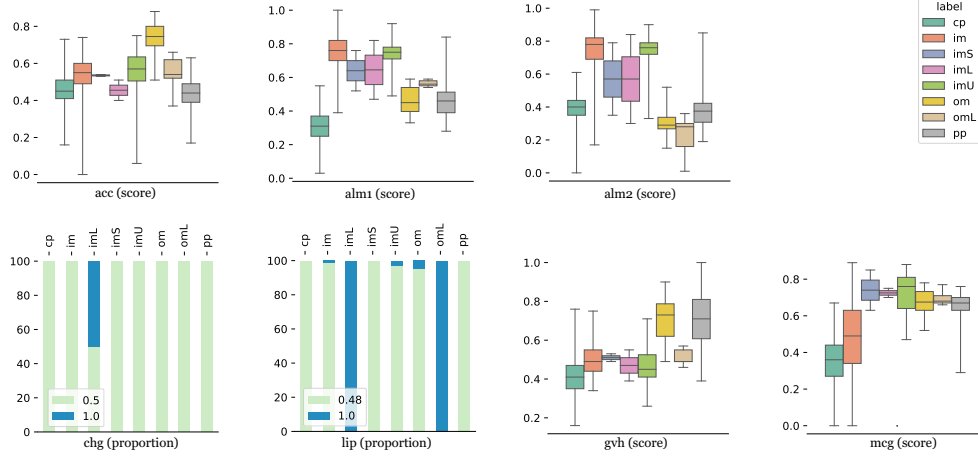


Figure 5.5: Distribution for ecoli protein location categories. The distribution of the measurements for each of the features grouped in eight categories: cytoplasm (cp), inner membrane without sequence (im), periplasm (pp), inner membrane with incleavable sequence (imU), outer membrane (om), outer lipoprotein (omL), inner lipoprotein (imL) and inner cleavable sequence (imS). Note that two of the features (chg and lip) are discrete. In these cases, the proportions have been presented.

5.4.2 Analysis of the CBR retrieval performance

This section compares the performance of different CBR configurations to retrieve and appropriately rank by similarity observations from the database. Firstly, the effect of a number of distance metrics was explored under equal feature weights (see Tables 5.9 and 5.10). After this, a total of 8008 different weights were evaluated using grid search (see Table 5.2). The top four CBR configurations were selected and included in Table 5.11. The metrics included from left to right are: normalized discounted cumulative gain (nDCG) score, Cohen’s kappa coefficient (KAPPA), sensitivity (SENS) and specificity (SPEC). For the sensitivity and specificity, micro/macro strategies to compute the scores on multiclass problems are presented. Since the CBR employs distance metrics, the features were normalized. The grey shade indicates for each scenario the model which provides the best performance.

Table 5.9: Comparison of distance metrics (I).

		nDCG	KAPPA	SENS	SPEC
UNI	MAN	0.787*	0.82	0.88/0.72	0.98/0.98
	EUC	0.799†	0.83	0.88/0.74	0.98/0.98
	TRI	0.798†	0.83	0.90/0.75	0.98/0.98
	CUA	0.798†	0.83	0.88/0.74	0.98/0.98
	QUI	0.798†	0.83	0.88/0.74	0.98/0.98

† and * indicate groups where difference is not significant.

Description: This table shows the performance of a CBR in a multiclass problem with eight imbalanced categories. The weights are equal for all the features (uniform). The number of neighbors investigated was k=1..20.

Table 5.10: Comparison of distance metrics (II).

		nDCG	KAPPA	SENS	SPEC
UNI	BRAY	0.780	0.79	0.85/0.72	0.98/0.97
	CAN	0.767	0.77	0.84/0.70	0.97/0.97
	CHEB	0.786*	0.83	0.89/0.73	0.98/0.98
	CORR	0.711‡	0.75	0.83/0.64	0.97/0.97
	COS	0.708‡	0.74	0.82/0.66	0.97/0.97

‡ and * indicate groups where difference is not significant.

Description: This table shows the performance of a CBR in a multiclass problem with eight imbalanced categories. The weights are equal for all the features (uniform). The number of neighbors investigated was k=1..20.

The effect of the distance metric is quite similar to the one explained for the iris flower dataset (see section 5.3); that is, homogeneity among minkowski-based distances and significantly worse performance for the CORR and COS metrics. The euclidean distance (EUC) is selected for further investigation. Conversely, in this experiment the difference in performance between UNI and SEARCH_{i=1..4} is not statistically significant. Note that the uniform weights configuration was ranked 18th. In fact, the feature weights are relatively uniform in the top four weight configurations identified through grid search.

Table 5.11: Comparison of weight configurations.

		nDCG	KAPP	SENS	SPEC	weights
EUC	UNI	0.799 [†]	0.83	0.88/0.74	0.98/0.98	[1/7, 1/7, 1/7, 1/7, 1/7, 1/7, 1/7, 1/7]
	SEARCH ₁	0.804 [‡]	0.84	0.89/0.77	0.98/0.98	[0.1, 0.2, 0.2, 0.1, 0.1, 0.2, 0.1]
	SEARCH ₂	0.804 [‡]	0.83	0.89/0.76	0.98/0.98	[0.1, 0.2, 0.1, 0.1, 0.2, 0.2, 0.1]
	SEARCH ₃	0.803 [†]	0.83	0.89/0.70	0.98/0.98	[0.1, 0.2, 0.0, 0.1, 0.3, 0.2, 0.1]
	SEARCH ₄	0.802 [†]	0.84	0.89/0.77	0.98/0.98	[0.1, 0.2, 0.2, 0.1, 0.2, 0.1, 0.1]

[†] indicates groups where differences are not statistically significant.

Description: This table shows the performance of a CBR in a multiclass problem with eight imbalanced categories. A total of 8008 weights were evaluated in grid search; the top four are presented. The uniform weights (UNI) are shown as reference. The number of neighbors investigated was $k=1..20$.

The consequences of class imbalance are visible in the sensitivity value. For instance, when the "micro" strategy is used to compute the sensitivity each individual decision is weighted equally. Thus, the sensitivity is biased to measure how effective the classifier performs on the large classes in the collection. On the contrary, in the "macro" strategy all the categories are weighted equally. Thus, the sensitivities of small classes, which are generally low, contribute equally to the overall sensitivity. As a result, the overall sensitivity decreases considerably. The issue of class imbalance considerably affects the performance of CBR systems. Note that in order to classify a new observation, CBR retrieves from the case-base those observations that are similar. After this, it decides the category through a "voting" strategy. In scenarios where the size of the classes differs considerably, observations from large classes will often outnumber those from small classes in the retrieved list leading to misclassification.

The left graph in Figure 5.6 presents the evolution of the NDCG score through different lengths of the ranking list (k) for two different weight configurations (UNI and SEARCH₁). These configurations present very similar performances. The right graph in Figure 5.6 presents the evolution of the NDCG score for each category independently. The best performance is obtained for *cp*, *om* and *pp* which have been evaluated on 40, 6 and 14 query observations respectively. The performance of *omL* also seems high but it is not reliable for two reasons: only one query observation was evaluated and the case-base contained only one *omL* observation. However, the *omL* observation from the case base was indeed retrieved and ranked high for all k . The opposite occurs in *imS* where the only sample existing in the database was never retrieved. For this reason, the performance for this class is the worse. It is also interesting the relation between *imU* and *im* which present a very similar behaviour. Their performances are average and the reason is probably overlapping between the two classes with cases being retrieved somehow indistinctively.

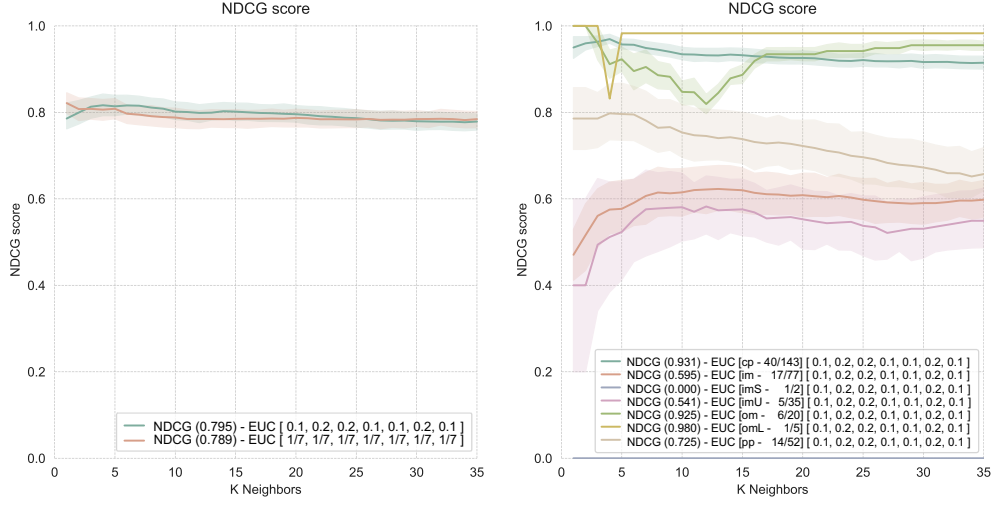


Figure 5.6: nDCG comparison On the left, evolution of the nDCG score through different k on two scenarios: (i) uniform weights (UNI) and (ii) best weight configuration found using grid search (SEARCH₁). On the right, the evolution of nDCG for each category independently. The information in the legend includes the name of the score (nDCG), the mean for all k , the distance metric, the number of query/-database cases and the feature weights.

To conclude, it is important to highlight that the number of weight combinations evaluated through grid search increases considerably from 286 (four features) to 8008 (seven features). As such, it is not feasible to perform grid search with a large number of features since the total number of possible weight combinations increase considerably. Moreover, note that grid search considers a set of discrete values for the weights while the domain is continuous. Thus, it would be interesting to identify an appropriate feature weight configuration through the use of optimization algorithms.

5.5 EXPERIMENT III: INTENSIVE CARE UNIT

The aim of this experiment is to evaluate the performance of the CBR algorithm with feature weights defined by the infection specialists. For such purpose, an internal prospective audit log of data from the CDSS was analysed. In order to evaluate the usability of the system and obtain further feedback, an electronic survey was sent to participants after the trial period. Moreover a further usability study with a wider range of health care workers including physicians and nurses was performed.

5.5.1 Pilot study by infection specialists

The pilot study was conducted by nine infection specialists during a four weeks period [35]. During this study, 20 ward rounds were undertaken with a total of 83 consultations to the CDSS. The CDSS was used for a total of 37 patients. The system correctly identified and pulled demographic information for all the patients. In addition, the system retrieved 3,898 biochemistry and haematology results. A total of 31 microbiology results were available for 14 of these patients as follows: 5 blood cultures, 5 wound cultures, 6 urine cultures, 7 sputum cultures and 8 screening cultures. Thus, the fidelity of the retrieval was 100%. Note that access was not provided to identify those cultures requested with a no-growth or pending outcomes. The antimicrobial therapy was adjusted on 31 episodes out of 83 consultations. In the majority of the cases (26/31, 84%) the most successful and similar cases identified by the CDSS advocated the same regimens as those recommended by the infection specialists. These regimens typically considered cefuroxime and metronidazole, tazocin or meropenem. For those scenarios in which the advice provided by the CDSS differed considerably from those recommended by the infection specialists (5/31, 16%) there were cases with epidemiological contact with other patients (3/5, 60%) which led to the empiric use of colistin. Overall, the study demonstrated that a functional CDSS enables high fidelity real-time patient-level data to be presented at the point of decision making.

From the responses provided by the infection specialists, positive aspects and areas in need of further improvement were identified. Firstly, all the participants highlighted the utility of the CDSS to access real-time patient data at the point of decision making. For instance, a participant commented that the microbiology data section worked really well since susceptibility tests that were reported from the laboratory that morning were already presented at the point of care (see C1 in Table 5.12). Moreover, the participant pointed that in other circumstances it would have been required to find a computer to check those results. Beyond data provision, participants also provided positive feedback on the results presented by the CBR to support knowledge transfer between registrars and consultants (see C2 in Table 5.12). Note that registrars move around between different hospitals. One of the main areas suggested for improvement was to reduce the complexity of data entry which should pull the information automatically from the servers (see C3 in Table 5.12). This concern was raised by more than half of the participants. Two of the participants also suggested the inclusion of personalisation of dosing strategies (see C4 in Table 5.12). The feeling among infection specialists was that the choice of antibiotic within the ICU is usually fairly simple (see C4 in Table 5.12) and the system could be more useful in secondary care.

Table 5.12: Pilot study in the ICU: free comments.

	Comment	Outcome
C1	“it worked really well in place of the [microbiology results] folder – it had all the bloods and the micro results, including new results from the lab that morning. You didn’t even need to have to go and find a computer and log in to check something” – Infection registrar, 39-year-old female	Positive
C2	“the panel on the right [where the CBR information is displayed] was useful, mainly for highlighting the top two or three regimes that others had used for similar patients. Most of what we use boils down to a handful of antibiotics, but when you’re moving around between hospitals, it’s great to get a feel of how similar cases are treated in this hospital by the consultants who have been here for ages” – Infection registrar, 33-year-old female	Positive
C3	“the first screen takes too long. In an ideal world the ICU team will of course tell you all the information you need to put in there, but they often don’t and you need to move back and forward across the screen as you are having a dialogue with them, so you can’t spend the minute or so entering data that the screen needs. Once you get beyond that, it’s fine, but there needs to be a mechanism to pull in more of those bits [parameters]” – Infection registrar, 39-year-old female	Negative
C4	“for us [infection specialists] choice of antibiotic is usually fairly simple in ICU, it’s actually harder outside of ICU, but I can see how the system would be [very] useful for the intensivists or that. What would make it more useful would be if it calculated doses, depending on BMI, whether they’re [haemo]filtered or not, etc” – Infection registrar, 33-year-old female.	Neutral

Further comments can be found on the drafted article [35].

5.5.2 Usability study of the CDSS

A survey to evaluate the usability of the decision support system interface in a variety of health care workers was performed. The System Usability Scale (SUS) [36] is composed of 10 statements to which participants indicate their agreement from 1 to 5, where 5 indicates strongly agree. This is commonly denoted as Likert scale. Predefined rules for positive and negative statements are used to obtain the SUS contributions. For odd items the SUS contribution is obtained by subtracting one to the user response. For even items the user response is subtracted from 5. The SUS contribution for each statement (see Table 5.13) is sum and multiplied by 2.5 to calculate the final SUS score. This score ranges from 0 to 100 where the following categories have been defined in [37, 38]: (i) score of 0–25: worst imaginable; (ii) score of 25–39: poor; (iii) score of 39–52: OK; (iv) score of 52–73: good; (v) score of 73–85: excellent; and (vi) score of 85–100: best imaginable. This survey is technology-agnostic, quick, provides a single score and is non proprietary. A free-text box was added for additional comments and suggestions.

The SUS survey was completed by 10 different participants (80% males) from 27 to 51 years old where technical training in the use of the system was not provided. The profile of those participants was infection specialist (two), clinician (three), nurse (four) and a pharmacist. The SUS contribution for each statement is presented in the right column in Table 5.13. Additionally, a variety of comments were provided by participants and have been synthesized in the following bullet points:

Table 5.13: The original SUS statements [36], average agreement and SUS contribution.

SUS statements	Avg. rating	SUS contribution
1. I think that I would like to use this system frequently.	2.8	1.8
2. I found the system unnecessarily complex.	1.4	3.6
3. I thought the system was easy to use.	2.0	1.0
4. I think that I would need the support of a technical person to be able to use this system.	1.6	3.4
5. I found that the various functions in this system were well integrated.	3.0	2.0
6. I thought that there was too much inconsistency in this system.	0.2	4.8
7. I would imagine that most people would learn to use this system very quickly.	2.8	1.8
8. I found the system very cumbersome to use.	2.2	2.8
9. I felt very confident using the system.	3.0	2.0
10. I needed to learn a lot of things before I could get going with this system.	0.8	4.2

- There was a common concern among experienced clinicians and infection specialists in the use junior doctors would make of such large amount of data displayed in the interface. The decision support system has potential to help training junior doctors and improve their prescription practices, but it needs to narrow the presented information providing specific guidance.
- Clinicians consider the user interface intuitive and helpful for patient long-term monitoring and management, however it might sometimes be a bit busy and time consuming. Additionally, it does not entirely fit with the workflow followed by clinicians to prescribe antibiotic therapies.
- Some participants suggested the possibility of recording further parameters not necessarily related with infections. Note that one of areas for improvement derived from the responses provided by infection specialists (see C3 in Table 5.12) was to reduce the number of parameters manually collected.

The usability measured through the SUS survey was 68.5 which is about average and shows potential margin for improvement. A similar strength of agreement was shown by participants for the third and eight statements which had an average rating of 2.0 and 2.2 respectively. The SUS contribution for each statement was 1.0 and 2.8 respectively, indicating that the system is usable but not necessarily easy. Note that some wording used by the original SUS was suggested to be poorly understood by participants affecting the results [37]. As an example, the sixth statement which contributed the most contains the wording “too much” which might be unclear. Additionally, users may not have a good sense of the required complexity of such systems since there are no commonly known competing solutions and the wording “unnecessarily” in the second statement might have led to a higher contribution of 3.6. Therefore, the final SUS score is possibly slightly less than the one presented.

5.6 DISCUSSION

Decision support systems are being exploited in several areas such as business or economics but their acceptance by clinical staff is obstructing its use in hospitals and other clinical environments [39]. Designing a CDSS based on usability principles and simply providing the clinical information does not guarantee acceptance [40]. Other factors as accessibility, availability, ease of use, time requirements and integration into the clinical workflow are important and need to be considered [41]. Taking previous knowledge into consideration, a CDSS to support prescription of antibiotic therapies and patient monitoring at POC exploiting the CBR methodology was implemented.

5.6.1 Benefits and limitations of CBR

The CBR methodology is the process of solving a new problem based on the solutions of past similar problems. This methodology aligns with the cognitive process followed by humans for problem solving [7]. For such reason, it has been widely used in a variety of domains including medicine [5, 6]. The first challenge while designing a CBR model is to decide the descriptive variables and the similarity metrics for comparison. A wealth of research has investigated and proposed a number of similarity metrics [8, 42] yet the most popular is the euclidean distance.

The issue of missing data

The presence of missing features in the data is a common problem in clinical domains which is often addressed by imputing either the overall mean or median. Furthermore, recent studies have proposed to input the mean or median using exclusively the surrounding neighbours [43]. An alternative strategy sees the missing data as valuable information which is included as an additional category [44]. Thus, this strategy assumes that patients with the same missing parameters are similar. Note that physicians record medical features relevant to treat the patient. Thus the 'no-recording' actually contains certain degree of similarity.

The issue of class imbalance

The issue of class imbalance considerably affects the performance of CBR systems; in particular for those relying on the K-Nearest Neighbours (KNN) algorithm to make predictions. Note that in order to classify a new observation, CBR retrieves from the case base those observations that are similar. After this, it decides the category through a "voting" strategy. In scenarios where the representation among categories differs considerably, observations from large classes will often outnumber those from small classes in the retrieved list leading to misclassification. A number of approaches might be considered to address this issue. The most straightforward strategy relies in under-sampling the majority classes. On the other hand, a number of strategies are implemented to oversample the minority classes. The random oversampling technique duplicates entries while the Synthetic Minority Oversample Technique (SMOTE) creates new synthetic observations. While these strategies are appropriate in classification problems where the output provided to users is the final category. The main focus of this research was information retrieval. As such, the mentioned sampling techniques were discarded since synthetic observations would be presented to physicians in the retrieved ranked list.

Towards a preference-guided search process

In order to select an appropriate CBR configuration (distance metric and feature weights) the normalized Discounted Cumulative Gain (nDCG) score has been used to measure the quality of the retrieved rankings. This score uses the graded relevance of the results retrieved and their position in the rank. In this research, the graded relevance is binary; that is, whether or not the category matches. However, there is also a sense of similarity between the solutions that could be used to define the graded relevance [45]. This preference-based approach would replace the current graded relevance from "solution y solves problem x " (binary similarity) to weaker information of the form " y is better than z as a solution for x " (solution similarity). Note that the binary similarity used in this research is the most restrictive. Thus introducing the preference-based similarity would leverage the previously presented results.

To conclude, note that CBR systems are mostly used for classification and their performance is evaluated using metrics such as accuracy, sensitivity or specificity. However, this methodology can be also seen in the context of information retrieval as a recommender system. In this context, the performance is measured by the utility of the retrieved elements based on the position in the produced ranking. This approach is less intrusive and respects the autonomy of the physician to decide the appropriate antimicrobial therapy.

5.6.2 Feedback provided from users

In many circumstances, as complicated cases, providers prefer to consult their colleges or more specialised clinicians as infection specialists. This consultation among different members of the clinical staff was facilitated by the CDSS. In addition, it was believed to enhance decision making and homogeneous collection of vital signs among clinicians resulting in better prescribing practices. Furthermore, re-entering patient data to generate advices is a deterrent to use [40]. Although vital signs were manually imputed, the integration of the system into EHR to automatically extract pathology and microbiology laboratory results was a clear facilitator.

From the trials performed by infection specialists and the feedback obtained from the surveys, it is possible to conclude that the CBR algorithm is able to mimic the prescription practices of users. However, that is not enough to promote change in antibiotic prescription practices. Initially, as a quick solution infection specialists were keen in creating an "ideal" case base; that is, a set of cases with optimal antibiotic therapies according to infection guidelines and expert prescriptions. Such optimal therapies would then be suggested by the decision support system to further users. Hence, the knowledge would be transferred from infection specialists to other clinical staff (e.g. nurses and clinicians). Unfortunately, this approach presents several drawbacks. Creating a complete case base that covers the whole spectrum of possibilities is nearly impossible and time consuming. In addition, it is necessary to thoroughly revise the cases that are added to the case base. Note that future therapies prescribed by clinicians and recorded in the system might not agree with the infection guidelines, reshaping the case base and therefore altering CBR recommendations.

5.6.3 *The infection management pathway*

After discussion with a multi-professional team including physicians, nurses, pharmacists and non-medical researchers, a study to map the pathway followed by clinicians to prescribe antimicrobial therapies was performed [12]. The reported infection management pathway was defined as a stepwise Bayesian model of estimating probabilities in which each step adds systematically information to allow optimisation of decisions on diagnosis and management of infection. Initially, clinicians estimate the risk of infection and attempt to localize its source by looking at patient's physiological parameters. Once clinicians construct a picture of the severity of the infection, whether or not to initiate antimicrobial therapy is decided. In this step, local microbiology guidance provided within hospitals was the most commonly cited factor. Finally, they review and refine the treatment accordingly. During this process, physicians often construct some sort of mental probabilities which are powerful drivers for prescribing.

There is still much to be done to make this system work in routine clinical practice. Measuring absolute usability using a single metric is very challenging since many external factors influence the results (e.g. technical training or accessible technology). The feedback provided greatly helped to identify areas for improvement. Moreover, the results of the SUS survey are a good starting point to assess the benefits of including new components to tackle the identified weaknesses.

5.6.4 *Limitations*

In this research, there was a lack of clinical data relating the patient medical record and the appropriate antimicrobial therapy. Thus, it was not possible to investigate the effect of different distance metrics and weight configurations. This limitation was tackled by using the euclidean distance which performed best in previous experiments and assigning feature weights based on infection specialists consensus. Similarly, the variables and terminology were adapted according to the targeted department. In the first phase of the development, the variables that were key drivers of antimicrobial therapy selection in augmented care were identified [10]. Later, these variables and the corresponding terminology were modified for further use in secondary care [11]. As an example, the parameter denoted as immunosuppression in augmented care was renamed as human immunodeficiency virus (HIV) in secondary care. In addition, most clinicians follow algorithms and established metrics in their practice which could be included as a new parameter within the system. For instance, scores such as the systemic inflammatory response syndrome (SIRS) or the sequential organ failure assessment (SOFA) could be computed using the previously collected vital signs and symptoms.

The doctors involved in the development and evaluation of the system were infection specialists and other clinicians within a UK-based teaching hospital. However, two areas in which the system has enormous potential are the implementation in low and middle-income countries with limited resources and the adaptability of the system to be used by non-infection specialists. With this in mind, a number of factors such as the background of the physician managing the patient, the years of experience and whether or not the therapy prescribed adheres with accepted local/regional/national or international guidelines leads undoubtedly to a markedly heterogeneous database. As such, an important limitation is the need of a gold standard to review and decide which cases will

be retained in the case base. Note that retaining cases which were not optimally managed could lead to poor clinical management and antimicrobial use. As such, it would be extremely useful to extend this research to adapt and pilot the system in these new settings to evaluate the benefits and address the corresponding limitations.

5.7 CONCLUSIONS

The CBR methodology incorporated in the CDSS retrieved antimicrobial therapies similar to those prescribed by expert clinicians in the majority of cases in the preliminary trial. Moreover, it was widely accepted as a tool for information retrieval and long-term patient monitoring and management. Three main areas of improvement were identified from the feedback provided by expert clinicians: provide specific guidance which is an important requirement for junior clinicians, the need to reduce the time required to perform a consultation within the CDSS and better integration with the clinical workflow. After further investigation, the reported infection management pathway was defined as a multi-step Bayesian-like approach [12] which inherently tackles most of the previously defined weaknesses. Hence, there is potential to combine both strategies (information retrieval using CBR and further multi-step Bayesian-like assistance) into one single decision support system to increase acceptability among clinicians.

BIBLIOGRAPHY

- [1] Hui Wang and Yiming Kevin Rong. "Case based reasoning method for computer aided welding fixture design". *Computer-Aided Design* 40.12 (2008), pp. 1121–1132.
- [2] Aijun Yan, Weixian Wang, Chunxiao Zhang, and Hui Zhao. "A fault prediction method that uses improved case-based reasoning to continuously predict the status of a shaft furnace". *Information Sciences* 259 (2014), pp. 269–281.
- [3] Chun-Ling Chuang. "Application of hybrid case-based reasoning for enhanced performance in bankruptcy prediction". *Information Sciences* 236 (2013), pp. 174–185.
- [4] Dong-Xiao Gu, Chang-Yong Liang, Isabelle Bichindaritz, Chun-Rong Zuo, and Jun Wang. "A case-based knowledge system for safety evaluation decision making of thermal power plants". *Knowledge-Based Systems* 26 (2012), pp. 185–195.
- [5] Shahina Begum, Mobyen Uddin Ahmed, Peter Funk, Ning Xiong, and Mia Folke. "Case-based reasoning systems in the health sciences: a survey of recent trends and developments". *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 41.4 (2011), pp. 421–434.
- [6] Rainer Schmidt and Lothar Gierl. "Case-based reasoning for antibiotics therapy advice: an investigation of retrieval algorithms and prototypes". *Artificial intelligence in Medicine* 23.2 (2001), pp. 171–186.
- [7] Agnar Aamodt and Enric Plaza. "Case-based reasoning: Foundational issues, methodological variations, and system approaches". *AI communications* 7.1 (1994), pp. 39–59.
- [8] Gavin Finnie and Zhaohao Sun. "Similarity and metrics in case-based reasoning". *International journal of intelligent systems* 17.3 (2002), pp. 273–287.
- [9] Zoe Y Zhuang, Leonid Churilov, Frada Burstein, and Ken Sikaris. "Combining data mining and case-based reasoning for intelligent decision support for pathology ordering by general practitioners". *European Journal of Operational Research* 195.3 (2009), pp. 662–675.
- [10] Luke Stephen Prockter Moore. "Rapid infection diagnostics in the context of augmented care: investigating their role in antimicrobial prescribing and bacterial resistance". *PhD dissertation* (2016).
- [11] Timothy Miles Rawson. "Personalised antimicrobial management in secondary care". *PhD dissertation* (2018).

-
- [12] Timothy Miles Rawson, Esmita Charani, Luke Stephen Prockter Moore, Bernard Hernandez, Enrique Castro-Sánchez, Pau Herrero, et al. "Mapping the decision pathways of acute infection management in secondary care among UK medical physicians: a qualitative study". *BMC medicine* 14.1 (2016), p. 208.
- [13] Alice Richardson, Simon Hawkins, Fariba Shadabi, Dhamendra Sharma, John Fulcher, and B Lidbury. "Enhanced laboratory diagnosis of human Chlamydia pneumoniae infection through pattern recognition derived from pathology database analysis" (2008).
- [14] Alice M Richardson and Brett A Lidbury. "Infection status outcome, machine learning method and virus type interact to affect the optimised prediction of hepatitis virus immunoassay results from routine pathology laboratory assays in unbalanced data". *BMC bioinformatics* 14.1 (2013), p. 206.
- [15] Masami Minemura, Kazuto Tajiri, and Yukihiro Shimizu. "Liver involvement in systemic infection". *World journal of hepatology* 6.9 (2014), p. 632.
- [16] Haim Shmueli, Silvio Pitlik, M Drucker, Z Samra, H Konisberger, and Leonard Leibovici. "Prediction of mortality in patients with bacteremia: the importance of pre-existing renal insufficiency". *Renal failure* 22.1 (2000), pp. 99–108.
- [17] Rafael Sierra, Jordi Rello, María Angeles Bailén, Encarnación Benítez, Antonio Gordillo, Cristobal León, et al. "C-reactive protein used as an early indicator of infection in patients with systemic inflammatory response syndrome". *Intensive care medicine* 30.11 (2004), pp. 2038–2045.
- [18] W Ray Kim, Steven L Flamm, Adrian M Di Bisceglie, and Henry C Bodenheimer. "Serum activity of alanine aminotransferase (ALT) as an indicator of health and disease". *Hepatology* 47.4 (2008), pp. 1363–1370.
- [19] Roberto Rabello Filho, Leonardo Lima Rocha, Thiago Domingos Corrêa, Camila Menezes Souza Pessoa, Giancarlo Colombo, and Murillo Santucci Cesar Assuncao. "Blood lactate levels cutoff and mortality prediction in sepsis—time for a reappraisal? A retrospective cohort study". *Shock (Augusta, Ga.)* 46.5 (2016), p. 480.
- [20] Stephen Trzeciak, R Phillip Dellinger, Michael E Chansky, Ryan C Arnold, Christa Schorr, Barry Milcarek, et al. "Serum lactate as a predictor of mortality in patients with infection". *Intensive care medicine* 33.6 (2007), pp. 970–977.
- [21] Philipp Schuetz, Werner Albrich, and Beat Mueller. "Procalcitonin for diagnosis of infection and guide to antibiotic decisions: past, present and future". *BMC medicine* 9.1 (2011), p. 107.
- [22] Eli J Finkelsztain, Daniel S Jones, Kevin C Ma, Maria A Pabón, Tatiana Delgado, Kiichi Nakahira, et al. "Comparison of qSOFA and SIRS for predicting adverse outcomes of patients with suspicion of sepsis outside the intensive care unit". *Critical care* 21.1 (2017), p. 73.
- [23] Eamon P Raith, Andrew A Udy, Michael Bailey, Steven McGloughlin, Christopher MacIsaac, Rinaldo Bellomo, et al. "Prognostic accuracy of the SOFA score, SIRS criteria, and qSOFA score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit". *Jama* 317.3 (2017), pp. 290–300.

- [24] Shih-Wei Lin and Shih-Chieh Chen. "Parameter tuning, feature selection and weight assignment of features for case-based reasoning by artificial immune system". *Applied Soft Computing* 11.8 (2011), pp. 5042–5052.
- [25] Selma Limam Mansar, Farhi Marir, and Hajo A Reijers. "Case-based reasoning as a technique for knowledge management in business process redesign". *Electronic Journal on Knowledge Management* 1.2 (2003), pp. 113–124.
- [26] Aijun Yan, Hongshan Shao, and Zhen Guo. "Weight optimization for case-based reasoning using membrane computing". *Information Sciences* 287 (2014), pp. 109–120.
- [27] Jacek Jarmulak, Susan Craw, and Ray Rowe. "Genetic algorithms to optimise CBR retrieval". *European Workshop on Advances in Case-Based Reasoning*. Springer. 2000, pp. 136–147.
- [28] Mohamed Bekkar, Hassiba Khelouane Djemaa, and Taklit Akrouf Alitouche. "Evaluation Measures for Models Assessment over Imbalanced Datasets". *Journal Of Information Engineering and Applications* 3.10 (2013).
- [29] Pierre Baldi, Søren Brunak, Yves Chauvin, Claus AF Andersen, and Henrik Nielsen. "Assessing the accuracy of prediction algorithms for classification: an overview". *Bioinformatics* 16.5 (2000), pp. 412–424.
- [30] M Tromp, Benno Lansdorp, CP Bleeker-Rovers, JM Klein Gunnewiek, BJ Kullberg, and P Pickkers. "Serial and panel analyses of biomarkers do not improve the prediction of bacteremia compared to one procalcitonin measurement". *Journal of infection* 65.4 (2012), pp. 292–301.
- [31] Seri Jeong, Yongjung Park, Yonggeun Cho, and Hyon-Suk Kim. "Diagnostic utilities of procalcitonin and C-reactive protein for the prediction of bacteremia determined by blood culture". *Clinica Chimica Acta* 413.21-22 (2012), pp. 1731–1736.
- [32] Joseph Guillén, Jiankun Liu, Margaret Furr, Tianyao Wang, Stephen Strong, Christopher C Moore, et al. "Predictive models for severe sepsis in adult ICU patients". *Systems and Information Engineering Design Symposium (SIEDS), 2015. IEEE.* 2015, pp. 182–187.
- [33] Haibo He and Eduardo A Garcia. "Learning from imbalanced data". *IEEE Transactions on knowledge and data engineering* 21.9 (2009), pp. 1263–1284.
- [34] Arie Ben-David. "Comparison of classification accuracy using Cohen's Weighted Kappa". *Expert Systems with Applications* 34.2 (2008), pp. 825–832.
- [35] Luke SP More, Bernard Hernandez, Timothy M Rawson, Pau Herrero, Alison H. Holmes, and Pantelis Georgiou. "Intelligent clinical decision support for antimicrobial prescribing in critical care" ().
- [36] John Brooke et al. "SUS-A quick and dirty usability scale". *Usability evaluation in industry* 189.194 (1996), pp. 4–7.
- [37] Aaron Bangor, Philip T Kortum, and James T Miller. "An empirical evaluation of the system usability scale". *Intl. Journal of Human-Computer Interaction* 24.6 (2008), pp. 574–594.
- [38] Aaron Bangor, Philip Kortum, and James Miller. "Determining what individual SUS scores mean: Adding an adjective rating scale". *Journal of usability studies* 4.3 (2009), pp. 114–123.

-
- [39] Kensaku Kawamoto, Caitlin A Houlihan, E Andrew Balas, and David F Lobach. "Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success". *BMJ* 330.7494 (2005), p. 765. eprint: <http://www.bmj.com/content/330/7494/765.full.pdf>.
 - [40] Annette Moxey, Jane Robertson, David Newby, Isla Hains, Margaret Williamson, and Sallie-Anne Pearson. "Computerized clinical decision support for prescribing: provision does not guarantee uptake". *Journal of the American Medical Informatics Association* 17.1 (2010), pp. 25–33.
 - [41] Rosy Tsopra, Jean-Philippe Jais, Alain Venot, and Catherine Duclos. "Comparison of two kinds of interface, based on guided navigation or usability principles, for improving the adoption of computerized decision support systems: application to the prescription of antibiotics". *Journal of the American Medical Informatics Association* 21.e1 (2014), e107–e116.
 - [42] Francesco Ricci and Paolo Avesani. "Learning a local similarity metric for case-based reasoning". *International Conference on Case-Based Reasoning*. Springer. 1995, pp. 301–312.
 - [43] Steven Bogaerts and David Leake. "Facilitating CBR for incompletely-described cases: distance metrics for partial problem descriptions". *European Conference on Case-Based Reasoning*. Springer. 2004, pp. 62–76.
 - [44] Janet Kolodner. *Case-based reasoning*. Morgan Kaufmann, 2014.
 - [45] Amira Abdel-Aziz, Marc Strickert, and Eyke Hüllermeier. "Learning solution similarity in preference-based CBR". *International Conference on Case-Based Reasoning*. Springer. 2014, pp. 17–31.

PROBABILISTIC INFERENCE

This chapter describes research towards the implementation of a decision support module to provide stepwise guidance in infection management. Firstly, the chapter provides a brief introduction to supervised machine learning methods for classification (section 6.1). After this, the implemented methodology to obtain the predictive models is explained; from data assembling (subsection 6.2.2) to model evaluation (subsection 6.2.3). In particular, this methodology has been used to undertake four experiments: prediction of a culture-positive outcome (section 6.4), prediction of the most plausible sites of infection (section 6.5), prediction of culture-positive outcome combining the independent predictions of the most plausible sites (section 6.6) and prediction of the Gram type of a bacteria (section 6.7). To conclude, findings are discussed (section 6.9) and the main conclusions summarised (section 6.10).

6.1 BRIEF INTRODUCTION TO SUPERVISED MACHINE LEARNING

Machine learning handles the design and implementation of algorithms that can learn and extract patterns from data. The essential desired property for these methods is the ability to generalize and consequently perform well on unseen data. In general, three main categories are devised according to the information available in the training dataset.

- In *supervised learning* the observations include the input features and the corresponding label or output. As such, these algorithms attempt to find a function that maps the input domain with the output. For instance, classification is a classic supervised learning task in which an unseen observation is assigned to one output category.
- In *unsupervised learning* the observations only include the input features and are therefore unlabelled. As such, these algorithms attempt to define a function that describes the structure of the data. Since labels are not provided, evaluation of these algorithms is not straightforward. For instance, clustering is a common unsupervised learning task in which properties and patterns are used to group observations in categories.
- In *semi-supervised learning* both labelled and unlabelled observations are used to train the algorithm. While the acquisition of unlabelled data is rather straightforward, labelled observations often require human intervention. As such, the majority of the research in this area has focused on combining unlabelled data with a small set of labelled data with a reported increase in performance.

In this chapter, supervised learning methods are considered to map the input parameters with the output domain or label [1]. In particular, a subset of supervised machine learning algorithms for binary classification which are able to provide probabilities as outcome have been selected. Although not all the classifiers provide probabilities inherently (e.g. support vector machines), additional algorithms exist to produce an estimation. An overview of the algorithms evaluated in this chapter is presented below.

Gaussian Naïve Bayes

The Gaussian Naïve Bayes (GNB) method is based on applying Bayes' theorem with the assumption of independence between every pair of features. The likelihood function for each feature is assumed to be Gaussian and despite this simplifying assumption, it has worked quite well in many real-world situations (e.g. spam filtering) [2]. Moreover, they require a small amount of training data to estimate the necessary parameters, are extremely fast compared to more sophisticated methods and the generated models can perform on-line updates.

Decision Tree Classifier

A decision tree classifier (DTC) is a simple algorithm for classifying observations based on recursive partitioning given an attribute value. These classifiers have been widely used in clinical domains due to its simplicity to interpret and understand. The time required to train decision trees on large datasets is reasonable. However, as a result of the greedy strategy applied, they present high variance, are often unstable and tend to over-fit. Moreover, these methods do not tend to work well if decision boundaries are smooth; that is, exists significant overlap between categories.

Random Forest Classifier

Random forest classifier (RFC) is an ensemble learning method for classification based on DTCs. It constructs various decision trees trained with different portions of the data and outputs the class that is the mode of all the classifiers. The voting system of these classifiers often correct the habit of over-fitting the training set present in decision trees.

Support Vector Machine

Support Vector Machine (SVM) uses a kernel function to transform the training samples to a new space with higher dimensionality [3]. The boundary found in the high dimensional space is the hyperplane which maximizes the distance between classes (i.e. maximum margin hyperplane) and can have a non linear shape in the original data space. It employs the principle of Structural Risk Minimization to generalize better than conventional machine learning methods which employ Empirical Risk Minimization [4]. Though SVMs do not directly provide probability estimates, they may be calculated in the binary case using Platt scaling; that is, logistic regression on the SVM's scores [5].

Artificial Neural Networks

Artificial neural networks (ANN) are based on a collection of nodes which are connected. Each connection can transmit a signal, typically a real number, from one node to another. In most common implementations, the output of each node is computed by some non linear function. The connections between nodes are called edges and have typically a weight that adjusts as learning progresses. The weight indicates the strength of the signal. The most common ANN structure consists on several layers, the input layer with as

many nodes as input features, the hidden layers and the output layer. Artificial neural networks are fitted using the backpropagation algorithm [6].

6.2 MATERIALS AND METHODS

The methodology followed in this chapter to create the predictive models has been described graphically in Figure 6.1. In the first step, those observations which are duplicated, erroneous or highly deviated are discarded. After this, the remaining data is divided into complete observations (F_6) and incomplete observations ($\{F_n\}_{n=1}^5$). The complete observations are used to generate the predictive models. The incomplete observations are used to evaluate the robustness of these models against missing data. Moreover, 25% of the observations in F_6 are isolated in the hold-out set (HOS) to further evaluate the translational utility. The remaining observations are used to generate and preliminarily evaluate the predictive models.

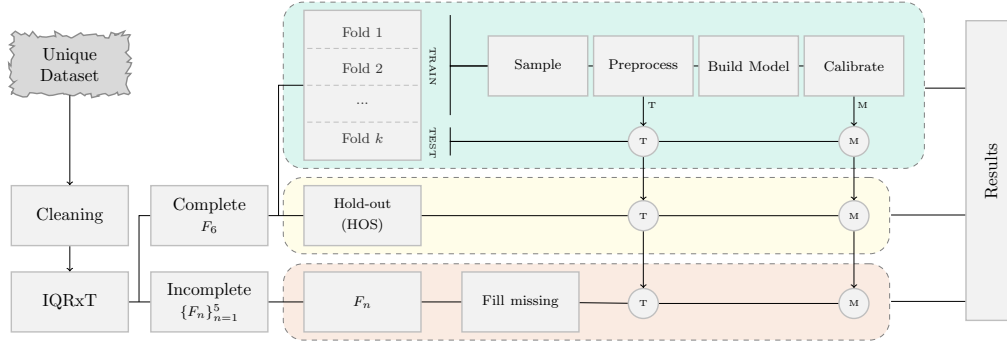


Figure 6.1: High-level methodology diagram for model creation and evaluation. First, data cleaning and outlier removal is performed. The remaining observations are grouped as complete or incomplete profiles. The former is further split into cross-validation Set (CVS) and hold-out set (HOS). Ten-Fold stratified cross-validation is performed on CVS and two outputs are obtained in this step: a preprocessing equation to transform new observations (T) and a calibrated model (M). It is important to highlight that sampling and preprocessing are performed using the train set while calibration is achieved from the test set. Finally, the performance of calibrated models is evaluated in HOS and $\{F_n\}_{n=1}^5$.

In this research, ten-fold stratified cross-validation has been used to assess how well the classifiers will generalize to an independent data set. This method divides the observations within the cross-validation set (CVS) in two categories: train and test. It is a common malpractice to perform data sampling on both categories. In this research, exclusively the train set is sampled, preprocessed and used to build the model as shown in Figure 6.1. From this procedure, a non-calibrated model and a preprocessing equation (T) to format unseen observations are obtained. Later, the model is calibrated using the observations within the test category. Finally, to assess the translational utility of these results into a clinical decision support system, the calibrated models are validated using HOS and $\{F_n\}_{n=1}^5$. Note that these observations were not seen during data sampling, data preprocessing, model training and model calibration.

6.2.1 Selected pathology biochemical markers

After reviewing the scientific literature and discussion of the qualitative investigations [7, 8] conducted by infectious disease experts (Luke Moore and Timothy Rawson), six routinely requested biochemical markers were selected (see Table 6.1). As explained in previous chapter, the aim of these biomarkers is to provide sufficient information to clinicians to evaluate the overall status of a patient and help identify underlying conditions such as organ dysfunction. For instance, creatinine is often requested to assess renal function and to determine how to alter antimicrobial doses. Note that these biomarkers were not selected with the specific purpose of identifying infections however, certain level of association between these biochemical markers and infectious diseases have been demonstrated in previous studies [9–14]. The inclusion of other biomarkers such as procalcitonin and lactate, which have been identified in the literature as important infection markers, was considered yet they were not routinely available for the majority of patients [15, 16].

Table 6.1: Description of selected laboratory biochemical markers.

Alanine aminotransferase	Alkaline phosphatase	Bilirubin
Alanine aminotransferase (ALT) is a transferase enzyme which is commonly used to screen liver problems. Moreover, significant high levels often suggest the existence of other medical problems such as viral hepatitis or infectious mononucleosis. The standard reference range is 10–40 ui/L with marked diurnal variations.	Alkaline phosphatase (ALP) is a protein enzyme which is commonly used to determine diagnosis such as hepatitis or osteomalacia. It also plays an integral role in metabolism within the liver and development within the skeleton. The standard reference range is 44–147 iu/L yet it varies with factors such as age, gender, pregnancy or blood type.	Bilirubin (BIL) is a yellow pigment that occurs when blood cells break down too fast and the liver is unable to metabolize it. For this reason, high levels might be related with impaired liver function or blockage in the biliary drainage. Moreover, an increase in the break down of blood cells can be caused by an autoimmune disease or an infection.
10–40 iu/L	44–147 iu/L	1.71–20.5 umol/L
Creatinine	C-Reactive protein	White blood cell
Creatinine (CRE) is a compound produced by metabolism of creatine which is excreted in the urine. Since kidneys maintain creatinine in the normal range, it is often associated with impaired kidney function or kidney disease. Moreover some studies suggest that CRE inhibit bacterial replication [17]. The standard reference range is 45–90 umol/L yet it relates with the muscle mass.	C-Reactive protein (CRP) is a protein found in blood plasma which is mainly used as a marker of inflammation. It is synthesized in the liver and therefore indicates liver failure. The rate of CRP production also increases with infection, trauma or allergic reactions. The standard reference range is 0.8–3.0 mg/L yet it increases with ageing.	White blood cells (WBC), also called leukocytes, are an important part of the immune system. Abnormal levels of blood cell counts may indicate that the immune system is fighting an infection. Low levels are also relevant as they might indicate that the immune system of the patient is diminished. The standard reference range is 4–11 10^9 /L in adults.
45–90 umol/L	0.8–3.0 mg/L	4–11 10^9 /L

Keys: iu/L=international units per litre; umol/L=micromoles per litre; mg/L=milligrams per litre;

6.2.2 Assembling data for inference

In hospitals, data is compartmentalized with many distinct measurements of patient health being stored separately. In this chapter, pathology and microbiology data for patients from all hospital wards at Imperial College Healthcare NHS Trust were extracted. In the absence of a single database linking pathology with microbiology data, these two different data sources were combined to create a unique dataset of daily profiles to perform probabilistic inference. Each profile (see Figure 6.2) has the daily symptoms of a patient represented by the selected biochemical markers (feature vector), the infection

condition extracted from the microbiology data (label) and additional information such as the patient identification number (PID) or date.

Unfortunately, the labels collected in these databases are recorded for purposes other than retrospective data analysis and it is difficult to define a ‘ground truth’. Initially, all daily profiles were labelled as culture-negative (C-). Then, any daily profile available for a patient with a positive culture in the microbiology data within two days difference was assigned to the culture-positive (C+) category. This assumption comes from antimicrobial susceptibility tests taking from 24h to 48h and antibiotics often needing a period of time to kill or stop bacterial growth. Note that assigning profiles to the culture-negative category by default clearly produces mislabelled data. To tackle this issue, profiles for those periods of time in with no susceptibility tests available were discarded. In addition, culture-negative profiles were removed if culture-positive profiles were present in a single patient admission.

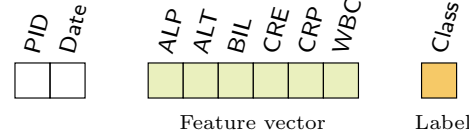


Figure 6.2: Biochemical markers: daily profile.

The profile has three main sections: metadata, feature vector and label. *PID*=patient identification number; *ALP*=alkaline phosphatase; *ALT*=alanine aminotransferase; *BIL*=bilirubin; *CRE*=creatinine; *WBC*=white blood cell; *CRP*=c-reactive protein;

6.2.3 Challenges in clinical data: preprocessing

In machine learning applications, data preprocessing is a common step that becomes critical when dealing with data obtained from clinical environments [18]. First, class imbalance must be tackled since unequal class distributions arise naturally. Also, data corruption is frequent [19] which can be classified as erroneous data, missing data and imprecise data. The steps followed in data preprocessing are briefly explained below.

Detection of outliers

The importance of outlier removal to develop robust predictive models has been demonstrated previously [20]. In our data, outliers are mainly caused by two main factors: susceptibility tests not requested or wrongly reported (human errors) and inaccurate microbiology results (diagnostic device errors or limitations). To identify and discard them, the inter-quartile range rule (IQRxT) is applied to each category independently where T represents the threshold parameter. A threshold of T=1.5 is widely accepted and T=3 is considered to discard only extreme outliers.

Imputing missing data

A large proportion of profiles are incomplete; that is, they do not have results for the six selected biomarkers. The notation F_n is used to define the fraction of data in which profiles have exactly n biomarkers. Exclusively complete profiles are manipulated to generate the predictive models while incomplete profiles ($\{F_n\}_{n=1}^5$) are used to evaluate the robustness of such models for different degrees of missing variables. For simplicity, the statistical measure preferred for imputation of missing values is the median.

Dealing with class imbalance

The issue of class imbalance has been addressed with three different strategies: under-sampling the majority class ($RAND_U$), over-sampling the minority class ($RAND_O$) and using Synthetic Minority Oversampling (SMOTE) [21] which blends both sampling methods to build classifiers with better performance. As previously mentioned, it is a common malpractice to perform data sampling on the whole data since it generates artificial observations (non-real data) which are eventually used for evaluation. Moreover, since random oversampling just duplicates entries, identical observations would be seen in model construction and evaluation defeating the whole purpose of cross-validation.

Data scaling and preprocessing

Data scaling is a requirement for many machine learning algorithms and can favourably affect model performance [22]. In this research two approaches have been considered: data normalization which scales individual features to have unit form and data standardization which transforms features so they are normally distributed; that is, zero mean and unit variance. It is a common malpractice to perform data scaling on the whole data since it often leads to over-fitting. Note that it leaks information from observations that will be later used for evaluation to the process of model construction. As mentioned before, completely unseen observations must be used to evaluate the model.

6.2.4 Model evaluation

There are many different metrics for assessing the performance of classifiers [23, 24]. For binary classifiers, most of them are based on four simple measures: the number of true positives (TP), the number of false positives (FP), the number of true negatives (TN) and the number of false negatives (FN). Sensitivity, specificity and overall accuracy are commonly used to demonstrate classifiers performance [25–27]. Note however, that accuracy might not be appropriate when the class sizes differ considerably [28]. For detailed information of classifiers, receiver operating characteristic (ROC) and precision-recall (PR) curves are often presented [29, 30]. The ROC curve is created by representing the true positive rate against the false positive rate for different threshold settings while the PR curve represents precision against recall. The area under such curves is commonly used for comparison. It is important to mention that precision is affected by class proportions, and hence PR is conditioned too. On the contrary, sensitivity, specificity and ROC are agnostic to class proportions. The definition and equations of previously mentioned metrics are shown in Table 6.2.

In addition to previous metrics, the use of positive and negative predictive values to describe the performance of a diagnostic test might be of higher utility in clinical practice [31]. The positive predictive value describes the proportion of positives which are truly positive. Similarly, the negative predictive value describes the proportion of negatives which are truly negative. These metrics are not intrinsic to the test and also depend on the prevalence of the disease. The data, as described previously, were not collected purposely and is composed of a number of diseases from which the ‘ground truth’ and therefore the corresponding prevalence was not available. For such reason, a good balance between sensitivity and specificity, which can be evaluated through the geometric mean, has been considered to select the best performing models.

Table 6.2: Evaluation metrics: description and equation.

Metric	Description	Equation
Sensitivity	Proportion of observed positives that are correctly identified as such (i.e. percentage of culture-positive profiles correctly identified as positive). Also called recall (REC) or true positive rate (TPR).	$SENS = \frac{TP}{TP+FN}$
Specificity	Proportion of observed negatives that are correctly identified as such (i.e. percentage of culture-negative profiles correctly identified as negative). Also called true negative rate (TNR).	$SPEC = \frac{TN}{TN+FP}$
ROC	This curve illustrates the performance of a binary classifier as its discrimination threshold is varied by plotting true positive rate (TPR) against false positive rate (FPR). It is related to cost/benefit analysis of diagnostic decision making.	
PR	This curve represents precision against recall where high scores for both shows that the classifier is returning accurate results (high precision) as well as returning a majority of all positive results (high recall).	

Keys: ROC=receiver operating characteristic curve; PR=precision recall curve; SENS=sensitivity; SPEC=specificity; TP=true positive; TN=true negative; FP=false positive; FN=false negative; TNR=true negative rate;

6.2.5 Model calibration

Properly calibrated classifiers provide a probability which incorporates an insight on the confidence of the prediction. For instance, in binary classification, among the samples to which a calibrated model gave a probability close to 0.8, approximately 80 % of samples actually belong to the positive class. Some models return well calibrated predictions by default while others introduce bias (e.g. GNB pushes probabilities to 0 or 1). They can be calibrated using a dataset not seen during training [32]. In this research a non-parametric approach based on isotonic regression was selected.

6.2.6 Statistical analysis

The statistical significance of the differences between the classifiers was determined using the non-parametric test (Kruskal-Wallis or one-way ANOVA on ranks) where the significance level was set at $p < 0.05$. Post-hoc analysis (Fisher's LSD) was used to determine pairwise differences.

6.2.7 Outline of the experiments conducted

This section briefly outlines the experiments further described in this chapter. For such purpose, the experiments have been depicted graphically in Figure 6.3. The first experiment (Experiment I in Figure 6.3) aims to discern between positive and negative microbiology cultures. After this, a second experiment (Experiment II in Figure 6.3) is conducted to assess whether it is possible or not to identify the source of infection. For this purpose, culture positive profiles are further divided into five sets according to the specimen collection site: blood, urine, sputum, wound and others. The third experiment (Experiment III in Figure 6.3) compares two different strategies to predict culture positivity. The first strategy uses the single classifier computed in the first experiment.

The second strategy evaluates an ensemble classifier which combines the best classifiers obtained in the second experiment. The next experiment assesses whether it is possible or not to distinguish between Gram-positive and Gram-negative bacteria. To conclude, the last experiment evaluates the robustness of the selected estimators in scenarios with missing data and imbalanced classes.

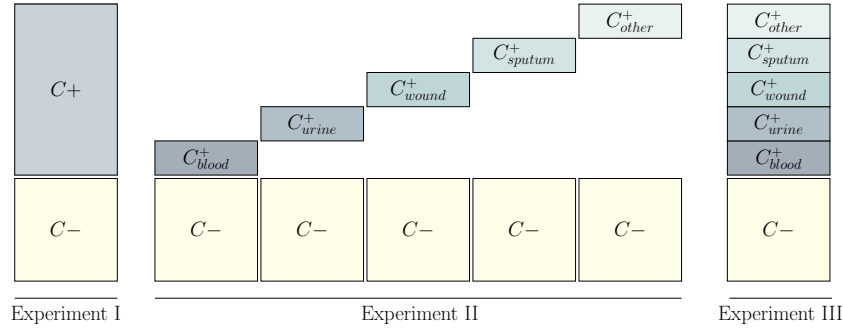


Figure 6.3: Graphical description of the experiments. The experiments described in this diagram are: (i) prediction of positive culture (ii) prediction of site of infection and (iii) ensemble method to predict positive cultures. The keys used are culture positive (C^+) and culture negative (C^-). Two additional experiments have been performed but are not included in the graph. One assesses whether it is possible or not to distinguish between Gram-positive and Gram-negative bacteria. The other evaluates the robustness of the selected estimators under two different conditions: missing data and imbalanced classes.

6.3 DATA ANALYSIS

The data used in this research was provided by the Imperial College Healthcare NHS Trust, which comprises three separate hospitals, totalling 1,500 beds and serving a population of 2.5 million citizens. The data comprises pathology laboratory results and microbiology data for 2014 and 2015. The combination of daily pathology laboratory results with the corresponding microbiology outcomes yielded over one and a half million profiles for more than half a million different patients. From these data, 43,497 (2.7%) profiles for 12,099 (2.1%) patients were assigned to the culture-positive category. Therefore, classes were clearly imbalanced with culture-negative constituting the majority.

6.3.1 Laboratory tests frequency

The number of laboratory tests requested per biomarker is explained for both categories (culture-negative and culture-positive) independently in Table 6.3. The notation F is used to categorize profiles according to the number of biomarkers available. Hence, F_2 contains all profiles with exactly two biomarkers. Obviously, some biomarkers are requested more frequently than others; from the instances presented in Table 6.3, the corresponding proportions are displayed for culture-negative (Figure 6.4a) and culture-positive (Figure 6.4b) categories. The most requested biomarkers are WBC and CRE for both categories. It is worth stressing that CRP is requested more frequently for infected patients. Its presence is almost double; from 11% in culture-negative profiles to 18% in

Table 6.3: Biochemical markers and daily profiles: frequency and completeness overview.

		ALP	ALT	BIL	CRE	CRP	WBC	All Tests	Profiles
C-	F ₁	10858	236	327	53443	10477	191213	266554	266554
	F ₂	11654	492	889	81337	25959	94605	214936	107468
	F ₃	51047	27921	28506	131058	113049	130870	482451	160817
	F ₄	135450	97665	101738	112962	36446	59607	543868	135967
	F ₅	412266	386171	409873	404555	58530	391120	2062515	412503
	F ₆	517397	517397	517397	517397	517397	517397	3104382	517397
	Total	1138672	1029882	1058730	1300752	761858	1384812	6674706	1600706
C+	F ₁	40	5	7	412	267	1445	2176	2176
	F ₂	103	12	20	1458	1140	1983	4716	2358
	F ₃	484	85	121	7671	7367	7621	23349	7783
	F ₄	2395	373	578	2308	1946	2096	9696	2424
	F ₅	5277	3043	5145	5165	3106	4674	26410	5282
	F ₆	23474	23474	23474	23474	23474	23474	140844	23474
	Total	31773	26992	29345	40488	37300	41293	207191	43497

Keys: C-=culture negative; C+=culture positive; F_n =set of profiles with n biomarkers available;

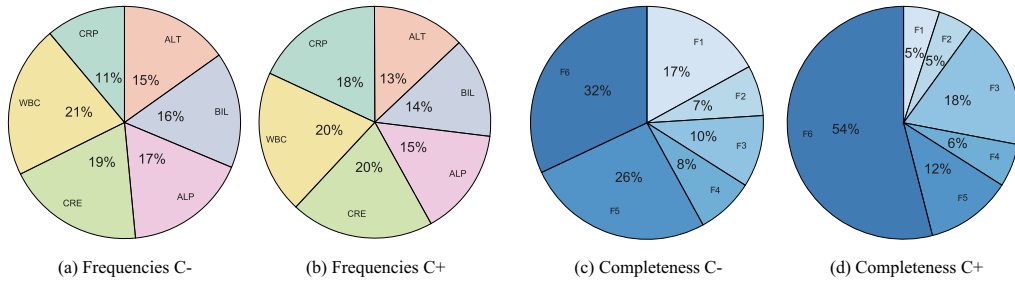


Figure 6.4: Biochemical markers: frequency and completeness. Percentages representing the frequency of each biomarker (a and b) and the completeness of profiles (c and d) for both culture-negative (C-) and culture-positive (C+) categories respectively.

culture-positive profiles. Although CRP would appear to be sufficient for infection detection by looking at the corresponding distribution (see Figure 6.5), it presents two main issues. First, it is the least requested of all biomarkers in the culture-negative category (11%), Second, it is an inflammatory marker which is not specific for infection. Finally, it does not provide any information regarding the location of the infection.

6.3.2 Profile completeness

A common problem in previous studies was missing data leading to incomplete profiles. Therefore, the proportion of profiles with different levels of completeness is displayed in Figures 6.4c and 6.4d. More than 50% of the culture-positive profiles are complete; that is, contain results for the six biomarkers. In contrast, the percentage of complete profiles drops to 28% for the culture-negative category. Note that evaluating four or more biomarkers are available ($\{F_n\}_{n=4}^6$) raises the percentages of profiles considered: culture-negative (65%) and culture-positive (71%). This is approximately two thirds of all available profiles. Thus, it is extremely important to generate classifiers that are able to provide accurate predictions in scenarios with missing biochemical markers to increase their use in clinical settings.

6.4 EXPERIMENT I: PREDICTION OF POSITIVE CULTURE

The aim of this experiment is to assess for a particular patient the likelihood of having a positive microbiology culture from six routinely collected pathology laboratory results. The requirement for culture is often independent of these variables as it is used to confirm infection and enable further susceptibility testing. Yet, predicting the likelihood of having a positive culture could potentially provide an insight on the potential utility of the requested test. Note that susceptibility testing takes between 24 and 48 hours.

6.4.1 Distributions of the selected biochemical markers

The density distribution for each biomarker is presented in Figure 6.5 for culture-positive and culture-negative categories. Most distributions are skewed (especially for C+) and robust measures for central tendency (median) and statistical dispersion (interquartile range) are used to describe them. Outliers were removed by applying the IQRx1.5 rule to both categories independently. The distance between medians for each category is clearly noticeable for CRP and appreciable to a lesser extent in WBC and ALP. On the other hand, there is no perceptible difference between the medians of culture-positive and culture-negative profiles for the rest of the biomarkers (ALT, BIL, CRE). Regarding statistical dispersion, CRP presents a huge contrast between the two categories, followed by CRE and ALP. Despite this contrast, the overlap in the IQR shows that individually these biomarkers do not provide information to differentiate between positive and negative categories.

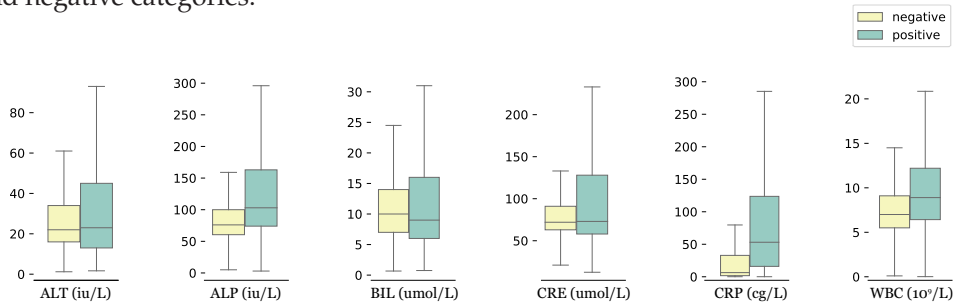


Figure 6.5: Biochemical markers: distributions for C- and C+. Distribution of measurements for each biomarker grouped in two categories: culture-negative (C-) and culture-positive (C+). The inter-quartile range rule with threshold of 1.5 (IQRx1.5) has been applied to each category independently to discard outliers.

6.4.2 Infection risk inference on complete profiles

This section compares the performance of different supervised learning models for the prediction of positive culture. The supervised learning models evaluated were GNB, DTC, RFC, SVM and ANN (see section 6.1). After evaluating several configurations for these models the best overall estimator was selected and included in the tables below. The presented metrics have been obtained using a balanced version of the hold-out set. Standardization performed consistently better than normalization and therefore only these results have been included. In order to deal with class imbalance, three sampling techniques (RAND_U, RAND_O and SMOTE) have been investigated. The metrics included from left to right are: area under the ROC curve (AUCROC), area under the PR curve (AUCPR where the subscript B indicates that classes were balanced), sensitivity

(SENS) and specificity (SPEC). The results presented in Table 6.4 considered the whole data. On the contrary, the results presented in Table 6.5 were obtained after discarding outliers using the IQR $\times$ 1.5 rule. The grey shade indicates for each scenario the model which provides the best performance.

Table 6.4: Comparison of sampling methods.

		AUCROC	AUCPR _B	SENS	SPEC
RAND _U	GNB	0.76	0.72	0.33	0.90
	DTC	0.79	0.76	0.77	0.68
	RFC	0.80	0.77	0.79	0.68
	SVM	0.79	0.74	0.76	0.69
	ANN	0.80	0.78	0.78	0.68
RAND _O	GNB	0.76	0.72	0.34	0.89
	DTC	0.75	0.73	0.63	0.74
	RFC	0.80	0.78	0.71	0.75
	SVM	0.80	0.74	0.77	0.69
	ANN	0.81	0.78	0.79	0.68
SMOTE	GNB	0.75	0.72	0.34	0.89
	DTC	0.73	0.70	0.30	0.89
	RFC	0.75	0.71	0.32	0.90
	SVM	0.79	0.73	0.75	0.70
	ANN	0.80	0.77	0.77	0.69

Description: This table shows the performance of a binary classifier in which the categories are negative culture (C-) and positive culture (C+). Three different sampling techniques (RAND_U, RAND_O and SMOTE) have been compared. Full dataset.

The performance of tree-based classifiers can be seen to vary according to the sampling technique used. It is particularly notable for the DTC classifier (see Table 6.4) which obtains quite high sensitivities when RAND_U (0.77) or RAND_O (0.63) are used to address the class imbalanced. Conversely, its sensitivity drops to 0.30 when SMOTE is applied. The tendency of DTC to over-fit might be the cause of such inconsistent behaviour. Note that DTC presents high sensitivities when the number of observations are reduced (RAND_U) or merely duplicated (RAND_O). Furthermore, such inconsistent behaviour remains, to a lesser extent, when outliers are discarded (see Table 6.5) with the following sensitivities: RAND_U (0.77), RAND_O (0.73) and SMOTE (0.65). The use of an ensemble approach (RFC) does not solve this inconsistent behaviour. On the other side, GNB, SVM and ANN are not affected by the sampling technique. However, GNB presents extremely low sensitivity values (0.33–0.34). The sensitivities for SVM (0.75–0.77) and ANN (0.77–0.79) are the highest among all the models. Moreover, they present a good balanced between sensitivity and specificity. These patterns also appear when outliers are discarded (see Table 6.5). As expected, the results obtained when outliers are discarded are superior than those using the whole data for all the models.

To summarise, the performance of the GNB, DCT and RFC models when outliers are not discarded (see Table 6.4) is poor. The sampling technique does not affect GNB yet it presents quite low sensitivities. In contrast, DTC is considerably affected by the sampling technique. SVM and ANN are quite consistent in all the scenarios and produce high sensitivity values. In further sections, only models generated using the SMOTE sampling technique are presented.

Table 6.5: Comparison of sampling methods (IQR $\times$ 1.5).

		AUCROC	AUCPR _B	SENS	SPEC
RAND _U	GNB	0.88	0.91	0.72	0.91
	DTC	0.90	0.92	0.77	0.88
	RFC	0.91	0.93	0.78	0.88
	SVM	0.89	0.91	0.77	0.87
	ANN	0.91	0.93	0.78	0.88
RAND _O	GNB	0.88	0.91	0.72	0.90
	DTC	0.88	0.91	0.73	0.88
	RFC	0.90	0.93	0.73	0.91
	SVM	0.89	0.91	0.77	0.87
	ANN	0.91	0.93	0.78	0.88
SMOTE	GNB	0.89	0.91	0.72	0.89
	DTC	0.88	0.91	0.65	0.95
	RFC	0.89	0.91	0.65	0.96
	SVM	0.89	0.91	0.77	0.87
	ANN	0.91	0.93	0.78	0.88

Description: This table shows the performance of a binary classifier in which the categories are negative culture (C-) and positive culture (C+). Three different sampling techniques (RAND_U, RAND_O and SMOTE) have been compared. Outliers discarded.

Table 6.6: Comparison inference of C+.

		AUCROC	AUCPR _B	SENS	SPEC
Positive	GNB	0.75	0.72	0.34	0.89
	DTC	0.73	0.70	0.30	0.89
	RFC	0.75	0.73	0.32	0.90
	SVM	0.79	0.74	0.75	0.70
	ANN	0.80	0.77	0.77	0.69

Description: This table shows the performance of a binary classifier in which the categories are negative culture (C-) and positive culture (C+). Full dataset.

Table 6.7: Comparison inference of C+ (IQRx1.5).

		AUCROC	AUCPR _B	SENS	SPEC
Positive	GNB	0.89	0.91	0.72	0.89
	DTC	0.88	0.91	0.65	0.95
	RFC	0.89	0.91	0.65	0.96
	SVM	0.89	0.91	0.77	0.87
	ANN	0.91	0.93	0.78	0.88

Description: This table shows the performance of a binary classifier in which the categories are negative culture (C-) and positive culture (C+). Outliers discarded.

The results presented in Table 6.6 considered the whole data. On the contrary, the results presented in Table 6.7 were obtained after discarding outliers using the IQRx1.5 rule. From the results presented in Table 6.7, the best estimators to predict culture positivity are: (i) GNB with priors of 0.5, since categories are balanced (ii) DTC with a minimum number of samples in a leaf of 50, a minimum number of observations in a node in order to be split of 200 and the gini impurity criterion (iii) RFC with 10 estimators (trees) with a minimum number of observations in a node of 100 (iv) SVM with penalty factor of $C=1.0$, a radial basis kernel where $\gamma=0.1$ and a tolerance of 0.001 and (v) ANN with a single hidden layer with 100 nodes, a ReLU activation function, alpha of 0.1, momentum of 0.9, learning rate of 0.001 and a tolerance of 0.0001. From the previously defined models, the highest levels of sensitivity correspond to SVM (0.77) and ANN (0.78). Moreover, both models also present an appropriate balance between sensitivity and specificity. Note that AUCROC (0.88–0.91) and AUCPR_B (0.91–0.93) are quite similar for all models. In order to further understand the behaviour of these models, the ROC curve (left), the PR curve (middle) and the calibration (right) are showed in Figure 6.6. The presented ROC curves show that the estimators have an appropriate trade-off between specificity and sensitivity. Moreover, from the PR curves it is possible to observe that they also exhibit a good trade-off between precision and recall. On the contrary, among the models considered, SVM and ANN also present the best calibration. Note that calibrated classifiers provide probabilities that can be interpreted as confidence intervals (see subsection 6.2.5).

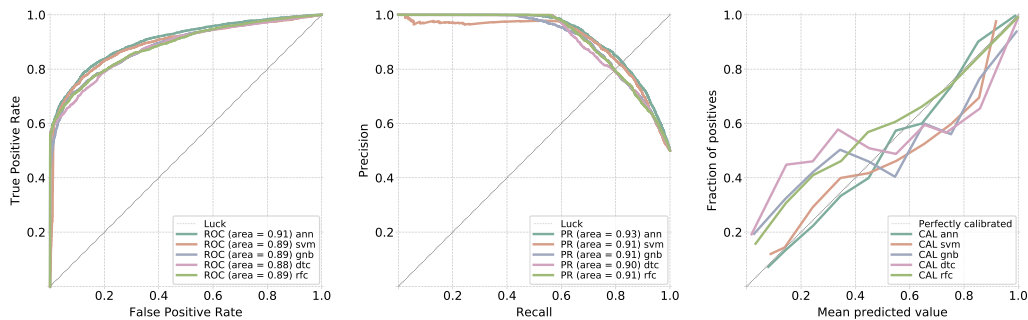


Figure 6.6: Positive culture inference: ROC, PR and calibration curves From left to right the graphs show the ROC, PR and calibration curves for those estimators constructed after discarding outliers using the IQRx1.5 rule. The scores have been computed on a balanced version of the hold-out set.

6.5 EXPERIMENT II: PREDICTION OF SITE OF CULTURE

The aim of this experiment is to assess whether it is possible or not to identify the source of infection. For this purpose, this experiment provides the likelihood of having a positive culture in a specific sample type. Note that specimen collection for susceptibility testing is conducted on a variety of sites such as the blood, the urine, the sputum or wounds. These sites are often closely related with the source of infection. As such, the models presented have to distinguish between negative cultures (C-) and positive cultures for one specific site (e.g. blood). In contrast to the previous experiment, focusing on infectious diagnosis for specific sites is quite common within the existing literature (see Appendix B.2).

6.5.1 Distribution of the selected biochemical markers

The density distribution for each biomarker is presented in Figure 6.7 for culture-negative (yellow) and culture positive within the blood, urine, sputum, wound and other cultures (shades of grey). In general, it is difficult to extract a common pattern within all the presented biochemical markers. The biomarkers can be divided into two groups by looking at the median of the distributions. One group of biomarkers (ALP, BIL and CRE) shows quite homogeneous medians for all the categories. In contrast, the other group (CRP, WBC and ALP) shows heterogeneous medians, especially for CRP followed to a lesser extent by WBC. Regarding the statistical dispersion of the distributions, blood usually presents the highest levels (ALT, BIL, CRP and WBC). As occurred in previous experiment, the overlap in the IQR shows that individually these biomarkers do not provide information to differentiate between positive and negative categories.

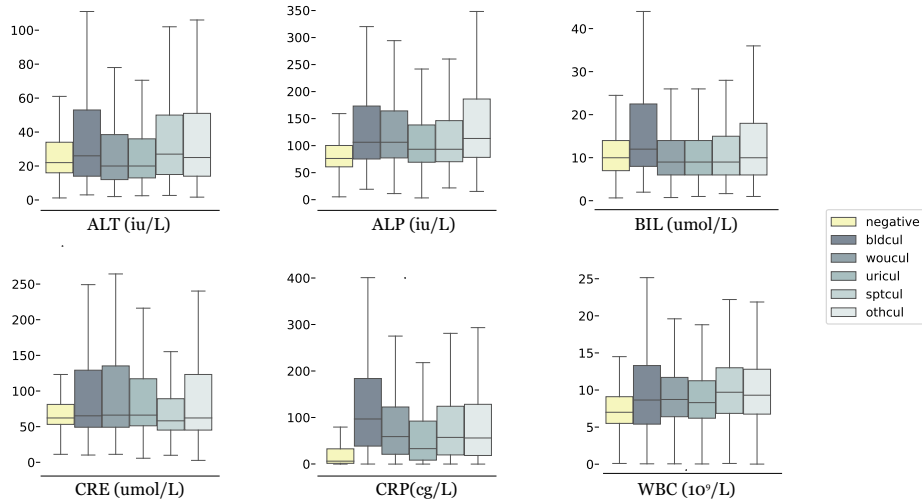


Figure 6.7: Biochemical markers: distributions for C- and sites. Distribution of measurements for each single biomarker grouped in six categories: culture-negative (C-) and culture-positive within the blood, urine, sputum, wound and other cultures. The inter-quartile range rule with threshold of 1.5 ($IQR \times 1.5$) has been applied to each category independently to discard outliers.

6.5.2 Site of infection risk inference on complete profiles

This section compares the performance of different supervised learning models for the prediction of positive culture in specific sites. The supervised learning models evaluated were GNB, DTC, RFC, SVM and ANN (see section 6.1). After evaluating several configurations for these models the best overall estimator was selected and included in the tables below. The results presented in Table 6.8 considered the whole data. On the contrary, the results presented in Table 6.9 were obtained after discarding outliers using the IQRx1.5 rule. The grey shade indicates for each scenario the model which provides the best performance.

In general, the performance of the models when outliers are not discarded (see Table 6.8) is quite poor with extremely low sensitivities (0.30–0.34). Note that AUCROCs are still high (0.73–0.75) but the levels of sensitivity and specificity are quite imbalanced. This issue has been previously seen in the literature (see Appendix B.2). In contrast, SVM-based and ANN-based models provide high and balanced levels of sensitivity (0.69–0.80) and specificity (0.66–0.75). Moreover, the results provided by ANN-based models are un-excelled in all scenarios. For such reason, they have been highlighted with a grey shade. On the other hand, it is no surprise that the results obtained when outliers are discarded are superior to those using the whole data (see Table 6.9).

Table 6.8: Comparison site of infection.

		AUCROC	AUCPR _B	SENS	SPEC
Blood	GNB	0.80	0.79	0.47	0.89
	DTC	0.81	0.82	0.49	0.92
	RFC	0.85	0.85	0.54	0.92
	SVM	0.88	0.85	0.79	0.81
	ANN	0.89	0.88	0.82	0.80
Urine	GNB	0.63	0.60	0.44	0.73
	DTC	0.62	0.60	0.17	0.92
	RFC	0.64	0.63	0.11	0.96
	SVM	0.70	0.67	0.64	0.65
	ANN	0.71	0.69	0.65	0.66
Sputum	GNB	0.78	0.72	0.30	0.90
	DTC	0.72	0.72	0.25	0.94
	RFC	0.74	0.74	0.25	0.95
	SVM	0.84	0.80	0.76	0.77
	ANN	0.85	0.80	0.82	0.75
Wound	GNB	0.71	0.67	0.41	0.83
	DTC	0.69	0.69	0.26	0.91
	RFC	0.74	0.73	0.26	0.94
	SVM	0.80	0.76	0.71	0.75
	ANN	0.81	0.78	0.73	0.74
Other	GNB	0.76	0.73	0.38	0.88
	DTC	0.72	0.70	0.31	0.90
	RFC	0.75	0.73	0.31	0.91
	SVM	0.80	0.75	0.78	0.69
	ANN	0.82	0.78	0.79	0.70

Description: This table shows the performance of a binary classifier in which one class represents negative cultures and the other class positive cultures in one particular site (e.g. blood).

Table 6.9: Comparison site of infection (IQRx1.5).

		AUCROC	AUCPR _B	SENS	SPEC
Blood	GNB	0.95	0.96	0.83	0.91
	DTC	0.94	0.96	0.82	0.95
	RFC	0.95	0.97	0.82	0.97
	SVM	0.96	0.97	0.86	0.93
	ANN	0.96	0.97	0.87	0.93
Urine	GNB	0.82	0.85	0.59	0.87
	DTC	0.76	0.82	0.47	0.95
	RFC	0.78	0.83	0.45	0.98
	SVM	0.83	0.85	0.64	0.86
	ANN	0.82	0.85	0.66	0.83
Sputum	GNB	0.90	0.92	0.73	0.90
	DTC	0.87	0.90	0.65	0.94
	RFC	0.90	0.92	0.65	0.96
	SVM	0.92	0.93	0.81	0.91
	ANN	0.93	0.94	0.83	0.87
Wound	GNB	0.90	0.92	0.73	0.92
	DTC	0.85	0.90	0.62	0.78
	RFC	0.88	0.91	0.62	0.98
	SVM	0.92	0.93	0.78	0.91
	ANN	0.92	0.93	0.79	0.80
Other	GNB	0.91	0.93	0.75	0.91
	DTC	0.89	0.92	0.70	0.95
	RFC	0.91	0.93	0.69	0.97
	SVM	0.93	0.94	0.78	0.92
	ANN	0.93	0.94	0.80	0.90

Description: This table shows the performance of a binary classifier in which one class represents negative cultures and the other class positive cultures in one particular site (e.g. blood).

In Table 6.9, the results can be grouped into three main categories according to the performance of the models. The first category corresponds to prediction of culture-positivity in blood samples; that is, bacteremia. Overall, the models show outstanding levels of AUCROC (0.94–0.96), sensitivity (0.82–0.87) and specificity (0.91–0.97). In the second category are sputum, wound and other samples. The models within this group show decent levels of AUCROC (0.85–0.93), sensitivity (0.62–0.83) and specificity (0.78–0.95). However, these are inferior to those obtained for blood samples. The last category corresponds to prediction of culture-positivity in urine samples which shows to be the most challenging. One reason could be the high incidence of asymptomatic bacteriuria which is very common with ageing and in hospitalised elderly patients. Note that DTC-based models exhibit generally low sensitivity values. In contrast, SVM-based and ANN-based models sacrifice specificity in favour of sensitivity which results in better overall predictions. The performance of these two methods is very similar (see Table 6.9). For further comparison, the ROC, PR and calibration curves have been presented for SVM (see Figure 6.8) and ANN (see Figure 6.9). In overall, the behaviour is equivalent though ANN shows a slightly superior calibration of probabilities than SVM.

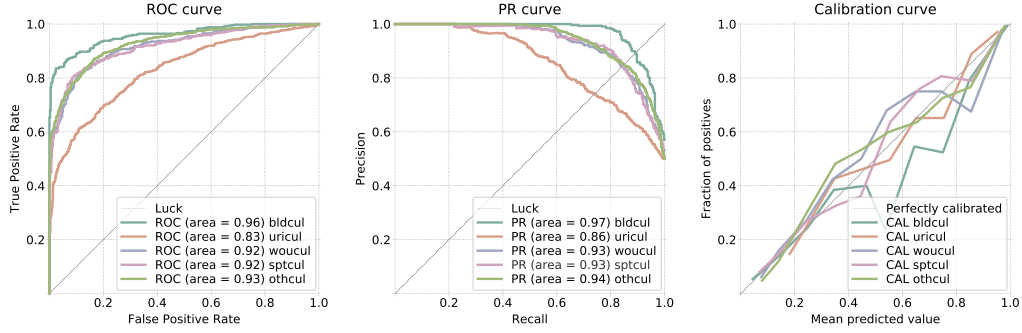


Figure 6.8: SVM-based estimators: ROC and PR curves (site). Inspection of the selected SVM-based estimators through ROC, PR and calibration curves for the different culture sites. The inter-quartile range rule with threshold of 1.5 ($IQR \times 1.5$) has been applied to each category independently to discard outliers.

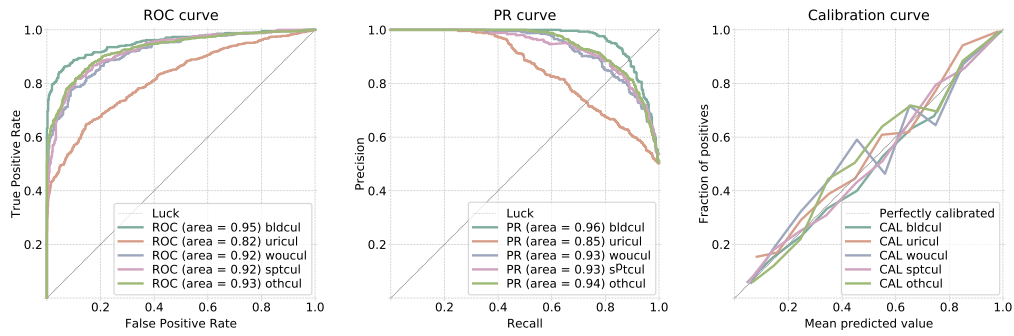


Figure 6.9: ANN-based estimators: ROC and PR curves (site). Inspection of the selected ANN-based estimators through ROC, PR and calibration curves for the different culture sites. The inter-quartile range rule with threshold of 1.5 ($IQR \times 1.5$) has been applied to each category independently to discard outliers.

To summarise, the most intelligible scenario is prediction of bacteremia. In contrast, identification of positive cultures within urine samples appears to be the most challenging. Note that existing research primarily focuses on bacteremia prediction while urine infections are studied to a lesser extent (see Appendix B.2). In this research, bacteremia prediction achieves the best scores partly due to the fact that all biochemical markers have been extracted from the blood. Hence, it would be interesting to incorporate other biochemical markers, such as those provided in urinalysis tests, to understand whether or not it is possible to boost the prediction of positive cultures in urine samples.

6.6 EXPERIMENT III: VOTING PREDICTION OF POSITIVE CULTURE

The aim of this experiment is to compare prediction of culture positivity using two different approaches. The first approach employs a single classifier to distinguish between negative cultures (C-) and positive cultures (C+) as described in section 6.4. In the second approach, an ensemble classifier to predict culture positivity is created by combining partial classifiers that predict site-specific positivity (see section 6.5). For this purpose the same training observations (CVS) and testing observations (HOS) have been used for both the single and the ensemble classifiers. The ensemble estimator is composed by five sub-estimators, one for each sample site. Thus, the ensemble estimator assigns a given observation to the positive class if at least one of the sub-estimators returned a positive. On the other hand, the probability is computed as the average of the sub-estimators probabilities. The results are presented below.

Table 6.10: Comparison inference of C+.

		AUCROC	AUCPR _B	SENS	SPEC
Positive single	GNB	0.73	0.69	0.30	0.89
	DTC	0.72	0.68	0.23	0.91
	RFC	0.72	0.68	0.24	0.90
	SVM	0.77	0.73	0.76	0.69
	ANN	0.78	0.75	0.77	0.68

Description: This table shows the performance of a binary classifier in which the categories are negative culture (C-) and positive culture (C+). No filter was applied.

Table 6.11: Comparison inference of C+ (ensemble).

		AUCROC	AUCPR _B	SENS	SPEC
Positive ensemble	GNB	0.73	0.70	0.42	0.83
	DTC	0.74	0.71	0.33	0.89
	RFC	0.74	0.71	0.33	0.88
	SVM	0.77	0.72	0.77	0.66
	ANN	0.78	0.74	0.77	0.67

Description: This table shows the performance of a binary classifier to distinguish positive and negative cultures by using a voting system that relies on site-specific classifiers. No filter was applied.

The ensemble estimators improve the sensitivity levels for those models which had previously low values. For instance, the sensitivity for GNB increases from 0.34 to 0.42. Similarly, the DTC increases from 0.23 to 0.33. This increase in sensitivity is due to the previously explained voting system; since now it is more common to return a positive class. However, it is not necessary to create an ensemble estimator for that purpose. Note that categories are assigned to the positive or negative category based on the obtained probability value and a threshold. Thus, decreasing the threshold would provide the same behaviour. On the other hand, ensemble estimators do not provide any further advantage to the already good performing SVM and ANN. In this research, the sub-estimators within the ensemble classifier share the same input features. However, the ensemble approach would perform better if each classifier has exclusively the input parameters which are required to predict culture positivity in a particular site.

6.7 EXPERIMENT IV: PREDICTION OF THE GRAM STATUS OF BACTERIA

The aim of this experiment is to assess whether it is possible or not to differentiate between Gram-positive and Gram-negative bacteria. As such, exclusively the culture positive dataset has been used. Firstly, models to distinguish between Gram-positive and Gram-negative bacteria using the whole culture positive dataset were created. After this, the culture positive dataset was further divided to determine whether it is possible or not to differentiate Gram status in specific sites. The results are presented in the tables below. Note that ANN-based models have been considered since they presented the best performance in previous experiments.

Table 6.12: ANN for Gram status prediction.

		AUCROC	AUCPR _B	SENS	SPEC
Positive	ANN	0.55	0.55	0.51	0.55
Blood	ANN	0.65	0.64	0.56	0.61
Urine	ANN	0.55	0.54	0.52	0.55
Sputum	ANN	0.63	0.63	0.56	0.64
Wound	ANN	0.59	0.58	0.57	0.55
Other	ANN	0.60	0.59	0.57	0.57

Description: This table shows the performance of a binary classifier in which one class represents Gram-negative bacteria and the other class Gram-positive bacteria. Whole data.

Table 6.13: ANN for Gram status prediction (IQRx1.5).

		AUCROC	AUCPR _B	SENS	SPEC
Positive	ANN	0.56	0.57	0.51	0.57
Blood	ANN	0.67	0.66	0.60	0.62
Urine	ANN	0.62	0.63	0.53	0.62
Sputum	ANN	0.66	0.70	0.58	0.66
Wound	ANN	0.65	0.63	0.67	0.51
Other	ANN	0.63	0.61	0.65	0.54

Description: This table shows the performance of a binary classifier in which one class represents Gram-negative bacteria and the other class Gram-positive bacteria. Outliers discarded.

In overall, biochemical markers do not contain enough information to appropriately distinguish between Gram-positive and Gram-negative bacteria. This is depicted by the results presented in Tables 6.12 and 6.13. As occurred in previous experiments, the highest scores are obtained when differentiating Gram-positive and Gram-negative bacteria in blood cultures which could be related with the fact that all biochemical markers selected are extracted from blood samples. However, the performance is clearly not sufficient to support antibiotic choice in clinical practice. Similar results are found for sputum with ROC, sensitivity and specificity of 0.66, 0.58 and 0.66 respectively. For the rest of datasets either sensitivity (in positive and urine) or specificity (in wound and other cultures) are near 0.5, which is equivalent to a random guess. As such, it is possible to conclude that prediction of Gram status bacteria is not achievable with the selected routinely collected biochemical markers [33].

6.8 EXPERIMENT V: MISSING DATA AND IMBALANCE CLASSES

The aim of this experiment is to understand the response of the selected models in real clinical settings. In particular two different scenarios have been considered: missing inputs and imbalanced class distributions. For such purpose, the behaviour of the selected ANN-based models is described in Figure 6.10. The estimators included in such figure are: (i) overall culture positivity (ii) culture positivity in blood (iii) culture positivity in urine (iv) culture positivity in sputum (v) culture positivity in wound and (vi) culture positivity in other sample types. The graphs used to describe the estimators are the ROC curve (first row), the PR curve (second row), the calibration curve (third row) and the thresholds curves (fourth row).

6.8.1 Inference on scenarios with missing data

This experiment describes the behaviour of the selected ANN-based models for different degrees of missing inputs. Since the models were trained on complete profiles (F_6) and the biomarker distributions are non-symmetrical (see Figure 6.5) the statistical measure preferred to input missing values is the median. Note that the median for each biomarker was extracted from the observations used to train the model. To study the behaviour of the models in scenarios with missing data, the profiles have been grouped in different categories according to the number of biochemical markers available ($\{F_n\}_{n=1}^6$). For instance F_2 contains all the profiles with two biochemical markers available (see Table 6.3). After this, the ROC curve has been computed for each of the groups (see first row in Figure 6.10).

The ROC curves obtained for $\{F_n\}_{n=4}^6$ are quite similar showing that all estimators perform without noticeable loss in performance if at least four biomarkers are available. The results obtained for F_3 are slightly inferior and the main drop in performance materializes for F_2 indicating insufficient information to perform inference. This is noticeable primarily in the sensitivity score.

6.8.2 Inference on scenarios with class imbalance

The estimators need to operate in scenarios where class imbalance is common. The ROC curve is a good indicator of overall performance but does not reflect the effect of class imbalance. For such reason the PR curve has been presented (see second row in Figure 6.10). The notation R80 indicates that 80% of observations belong to the culture-negative category. In scenarios with balanced classes (R50) the predictive model shows a good balance between precision and recall. The models are robust against high levels of class imbalance and the drop in performance occurs for scenarios with an imbalance ratio of approximately 1/9 (90%) or higher.

6.8.3 Probability calibration and thresholds

Properly calibrated classifiers provide a probability which can be directly interpreted as a confidence interval. For instance, in binary classification, among the samples to which a calibrated model gave a probability close to 0.8, approximately 80% actually belong to the positive class. The calibration of the selected models is shown in the third column in Figure 6.10. In overall, the calibration is excellent. Note that there are minor inconsistent variations in the categories wound and other. The observations are later assigned to one of the categories (either positive or negative) based on these probabilities. The most common strategy relies on defining a cut-off value which is denoted as threshold. Then, those observations with a probability below the threshold are assigned to the negative category. The default threshold value is often 0.5. The effect that the variation of the threshold has on the sensitivity and specificity of the model are shown in the threshold curves (see fourth row in Figure 6.10). The threshold values for which sensitivity and specificity are perfectly balance are within the range 0.4–0.5.

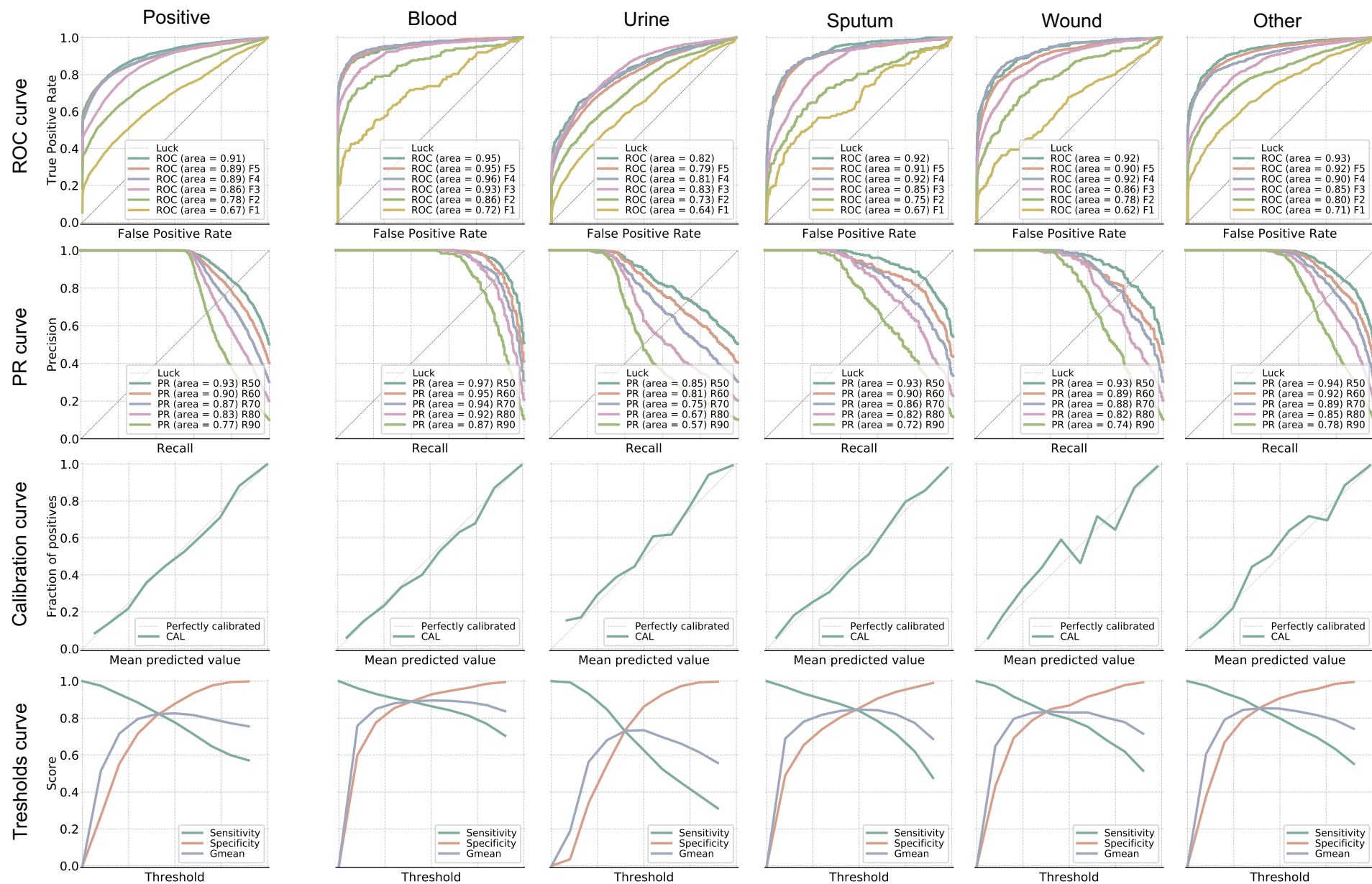


Figure 6.10: Summary of estimators: ROC, PR, calibration and threshold curves. The behaviour of the selected estimators to predict positive culture and the corresponding sites (blood, urine, sputum, wound and other). The results presented correspond to ANN-based estimators. The axes of the graphs are within the range [0,1]. *PR*=precision-recall; *ROC*=receiver operating characteristic; *CAL*=calibration.

6.9 DISCUSSION

In infectious diseases, antibiotic selection has been the main focus of clinical decision support systems (CDSSs) [34]. However, improving antibiotic selection does not necessarily imply a reduction in antibiotic prescription, it might even encourage it. Therefore, assisting clinicians by providing the risk of infection and most plausible sites for an individual patient can help discerning whether or not to initiate antibiotic therapy. This could potentially reduce the misuse of antibiotics. The main reasons obstructing inclusion in CDSSs were: (i) studies were highly specific by tackling individual microbes and single infections (sepsis is the most common) (ii) they required a high number of variables whose collection is laborious (iii) scenarios with missing data, which are very common in clinical environments, were completely ignored (iv) there was a lack of thorough evaluation of the models to understand their behaviour and support confidence.

6.9.1 *Selecting suitable biochemical markers*

The first challenge while designing a model for classification is deciding which input parameters are to be considered. After reviewing the scientific literature and discussion of the qualitative investigations [7, 8] conducted by infectious disease experts, six routinely requested biochemical markers were selected. The aim of these biomarkers was to provide sufficient information to clinicians to evaluate the overall status of a patient and help identify underlying conditions such as organ dysfunction. As such, these biomarkers were not selected with the specific purpose of identifying infections. Note that certain level of association has been demonstrated in previous studies [9–14]. The inclusion of other biomarkers such as procalcitonin and lactate, which have been identified in the literature as important infection markers, could greatly improve the presented results and their integration into clinical practice [15, 16].

6.9.2 *Addressing class imbalance*

The issue of class imbalance is common in detection and classification within health domains. For this reason, a number of sampling techniques were explored. Simple methods such as random under-sampling and/or over-sampling have proven to be valid in other domains. Undoubtedly, choosing an adequate sampling technique depends on the data, but it is clear that under-sampling potentially discards useful information and over-sampling replicates observations which might lead to over-fitting. Thus, the Synthetic Minority Oversampling Technique (SMOTE) was selected in this research.

6.9.3 *Effect of missing inputs in prediction*

Unfortunately, missing variables is a common problem in clinical data. Since this is a retrospective study, we have to deal with the fact that the data were not collected to generate a predictive model. For these reasons, it is highly desirable for a classification system to be robust to incomplete inputs. Both SVM and ANN classifiers are robust and operate without noticeable loss in performance if at least four biomarkers are present. DTCs are widely used in clinical research and the results obtained in this paper outperform those presented in similar studies [12, 13]. However, this method is the most affected by missing biomarkers as a result of the greedy strategy applied. In previous

studies RFC was selected as an ensemble method based on DTCs [12, 13, 35] to tackle this issue. The unexpected increase in sensitivity presented by DTC for scenarios with missing data was corrected. However, performance was found to be similar.

At the moment, the method selected to input missing values is the median. However, it would be incredibly useful to estimate the values of missing biochemical markers from those which are available. This could potentially boost the performance of the classifiers when biochemical markers are missing. Moreover, estimating the values of those biochemical markers that were not requested would help clinicians to better construct a clinical picture of the patient to assess the severity of the infection.

6.9.4 Selecting a suitable algorithm

From the obtained results, infection inference is feasible by using exclusively the six selected biomarkers with an AUCROC of approximately 0.8. In addition, sensitivity and specificity were both high and balanced in comparison to previous studies [35]. Despite the simplicity of Gaussian Naïve Bayes (GNB), it shows to be a good starting point which provides decent results. It also has additional desirable properties, namely that it requires a small amount of training data, it is computationally efficient and performs on-line updates. However, the best performance corresponds to (i) SVM classifier with penalty factor of $C = 0.1$ and radial basis kernel where $\gamma = 1.0$. and (ii) ANN classifier with 1 hidden layer with ten nodes and a ReLU activation function. The main disadvantage of SVM is the large amount of computational resources (memory and time) required. Conversely, ANN shows excellent performance with a drastic reduction in computational resources. The evaluation was performed on real observations (not synthetically generated) which were completely unseen during sampling, preprocessing, training and calibration. The latter is often ignored but necessary to guarantee that probabilities use the whole spectrum [0,1] and are informative by providing the degree of confidence in the prediction. The Bootstrap aggregating technique was explored to build ensemble classifiers based on GNB and SVM, but it did not provide any significant improvement. A summary of the properties of the algorithms used in this research is presented in Table 6.14.

Table 6.14: Algorithms: summary of properties

	GNB	DTC	RFC	SVM	ANN
Affected by sampling	No	Yes	Yes	No	No
Complexity of the algorithm	Low	Low	Medium	High	High
Performance of the algorithm	Poor	Poor	Medium	High	High
Number of configuration parameters	1	11	13	9	9
Sensitivity (true positive rate)	Low	High	Medium	High	High
Balance between sensitivity/specificity	No	No	No	Yes	Yes
Probability calibration	Poor	Poor	Poor	Good	Excellent
Time for training	Low	medium	High	High	Medium

Note: The number of configuration parameters is defined by the implementation of such algorithms provided in [36].

6.9.5 Translational utility

It is important to recognize that the evaluation in the training phase is different from the evaluation of the final model. The first phase is to tune the models' hyper-parameters and select the most effective and robust model during training. The second phase is to evaluate the final model after the training. Ideally, the test data of this phase reflects the class distributions of the original population even though such distributions are usually unknown. Since the SVM and ANN classifiers presented a robust response (and the highest sensitivities) in scenarios with missing data and imbalanced categories, they have been selected for further inclusion in the EPiC IMPOC (Enhanced Personalized and Integrated Care for Infection Management at Point of Care) decision support system to assist clinicians [37].

6.9.6 Limitations

In this research, profiles were assigned to the culture-positive (C+) category based on evidence of organism growth in the microbiology samples. Since there was a lack of no-growth evidence, remaining profiles were assigned to the culture-negative (C-) category. This limitation was tackled through data cleaning and outlier detection. However, providing no-growth evidence could boost performance even further. In addition, there was no additional information such as vital signs, symptoms or the final diagnosis. Due to this lack of additional data, the interquartile range (IQR) rule was applied to clean the data. Note that the most common strategies for data cleaning rely on demographics, symptoms, the final diagnosis and other related factors. The use of this information to clean the data could have helped to provide more accurate and informative predictions.

An important limitation in this research was the use of generic biochemical markers (such as c-reactive protein or creatinine) which were selected to monitor the patient and evaluate possible underlying conditions such as organ dysfunction. In order to boost the predictive performance of the algorithms, a more targeted set of biochemical markers could have been used. For instance, procalcitonin [16, 25, 26, 38, 39] and lactate [40, 41] have been recognised in the literature as important infection markers. Moreover, the parameters used to predict the likelihood of infection should be specific for each site. For instance, the introduction of biochemical markers extracted from urinalysis tests could help identifying asymptomatic bacteriuria and therefore boost the prediction of positive cultures in urine samples. The SVM-based and ANN-based models were thoroughly evaluated by replicating a wide range of conditions (some of them extreme) in which the classifier would operate. However, the prevalence of the disease was still unknown. Thus, an empirical study to quantify the prevalence and the costs of different mistakes (false positives and false negatives) would greatly help to understand their consequences and effects on prescription practices.

6.10 CONCLUSIONS

During clinical practice clinicians often construct some sort of mental probabilities which are powerful drivers for antimicrobial prescribing. This chapter demonstrates that the prediction of positive culture from six routinely collected biochemical markers is feasible (AUCROC=0.91; SENS=0.78; SPEC=0.88). Moreover, the identification of culture positiv-

ity in blood (bacteremia) obtained the highest performance (AUCROC=0.96; SENS=0.87; SPEC=0.93). On the other hand, culture positivity in urine samples is still challenging with very poor sensitivity values (AUCROC=0.82; SENS=0.66; SPEC=0.83) that could be improved through the introduction of more specific parameters such as those obtained from urinalysis. Moreover, other vital signs which are easy to collect such as heart rate or respiratory rate could also help identifying respiratory tract infections. The selected biochemical markers are not enough to predict the Gram status of a bacteria (see section 6.7) [33]. As such, it becomes essential to promote the development of point of care diagnostic tests to better identify the type of infectious pathogen [42] or its concentration [43]. The integration of such diagnostic tests within the CDSS would facilitate access to these results as soon as they become available promoting timely and optimal antimicrobial therapy selection.

BIBLIOGRAPHY

- [1] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. England: MIT press, 2012.
- [2] Vangelis Metsis, Ion Androutsopoulos, and Georgios Paliouras. "Spam filtering with naive bayes-which naive bayes?" *CEAS*. Vol. 17. 2006, pp. 28–69.
- [3] Bernard Hernández Pérez. *Multi-View Object Recognition and Classification. Graph-Based Representation of Visual Features and Structured Learning and Prediction*. 2013.
- [4] Hyunjung Shin and Sungzoon Cho. "How to deal with large dataset, class imbalance and binary output in svm based response model". *Proceedings of the Korean Data Mining Conference*. 2003, pp. 93–107.
- [5] John Platt et al. "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods". *Advances in large margin classifiers* 10.3 (1999), pp. 61–74.
- [6] Corinna Cortes, Xavi Gonzalvo, Vitaly Kuznetsov, Mehryar Mohri, and Scott Yang. "Adanet: Adaptive structural learning of artificial neural networks". *arXiv preprint arXiv:1607.01097* (2016).
- [7] Luke Stephen Prockter Moore. "Rapid infection diagnostics in the context of augmented care: investigating their role in antimicrobial prescribing and bacterial resistance". *PhD dissertation* (2016).
- [8] Timothy Miles Rawson. "Personalised antimicrobial management in secondary care". *PhD dissertation* (2018).
- [9] W. Ray Kim, Steven L. Flamm, Adrian M. Di Bisceglie, and Henry C. Bodenheimer. "Serum activity of alanine aminotransferase (ALT) as an indicator of health and disease". *Hepatology* 47.4 (2008), pp. 1363–1370.
- [10] Rafael Sierra, Jordi Rello, María Angeles Bailén, Encarnación Benítez, Antonio Gordillo, Cristobal León, et al. "C-reactive protein used as an early indicator of infection in patients with systemic inflammatory response syndrome". *Intensive Care Medicine* 30.11 (2004), pp. 2038–2045.
- [11] Masami Minemura, Kazuto Tajiri, and Yukihiro Shimizu. "Liver involvement in systemic infection". *World journal of hepatology* 6.9 (2014), p. 632.
- [12] Alice Richardson, Simon Hawkins, Fariba Shadabi, Dhamendra Sharma, John Fulcher, and B Lidbury. "Enhanced laboratory diagnosis of human Chlamydia pneumoniae infection through pattern recognition derived from pathology database analysis" (2008).

- [13] Alice M Richardson and Brett A Lidbury. "Infection status outcome, machine learning method and virus type interact to affect the optimised prediction of hepatitis virus immunoassay results from routine pathology laboratory assays in unbalanced data". *BMC bioinformatics* 14.1 (2013), p. 206.
- [14] W Ray Kim, Steven L Flamm, Adrian M Di Bisceglie, and Henry C Bodenheimer. "Serum activity of alanine aminotransferase (ALT) as an indicator of health and disease". *Hepatology* 47.4 (2008), pp. 1363–1370.
- [15] Roberto Rabello Filho, Leonardo Lima Rocha, Thiago Domingos Corrêa, Camila Menezes Souza Pessoa, Giancarlo Colombo, and Murillo Santucci Cesar Assuncao. "Blood lactate levels cutoff and mortality prediction in sepsis—time for a reappraisal? A retrospective cohort study". *Shock (Augusta, Ga.)* 46.5 (2016), p. 480.
- [16] Stephen Trzeciak, R Phillip Dellinger, Michael E Chansky, Ryan C Arnold, Christa Schorr, Barry Milcarek, et al. "Serum lactate as a predictor of mortality in patients with infection". *Intensive care medicine* 33.6 (2007), pp. 970–977.
- [17] Thomas McDonald, Kristen M Drescher, Annika Weber, and Steven Tracy. "Creatinine inhibits bacterial replication". *The Journal of antibiotics* 65.3 (2012), p. 153.
- [18] Ya-Han Hu, Wei-Chao Lin, Chih-Fong Tsai, Shih-Wen Ke, and Chih-Wen Chen. "An efficient data preprocessing approach for large scale medical data mining". *Technology and Health Care* 23.2 (Jan. 2015), pp. 153–160.
- [19] Alistair EW Johnson, Mohammad M Ghassemi, Shamim Nemati, Katherine E Niehaus, David A Clifton, and Gari D Clifford. "Machine learning and decision support in critical care". *Proceedings of the IEEE* 104.2 (2016), pp. 444–466.
- [20] Jason W Osborne and Amy Overbay. "The power of outliers (and why researchers should always check for them)". *Practical assessment, research & evaluation* 9.6 (2004), pp. 1–12.
- [21] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. "SMOTE: synthetic minority over-sampling technique". *Journal of artificial intelligence research* 16 (2002), pp. 321–357.
- [22] Lina Zhou, Shimei Pan, Jianwu Wang, and Athanasios V Vasilakos. "Machine learning on big data: Opportunities and challenges". *Neurocomputing* 237 (2017), pp. 350–361.
- [23] Mohamed Bekkar, Hassiba Khelouane Djemaa, and Taklit Akrouf Alitouche. "Evaluation Measures for Models Assessment over Imbalanced Datasets". *Journal Of Information Engineering and Applications* 3.10 (2013).
- [24] Pierre Baldi, Søren Brunak, Yves Chauvin, Claus AF Andersen, and Henrik Nielsen. "Assessing the accuracy of prediction algorithms for classification: an overview". *Bioinformatics* 16.5 (2000), pp. 412–424.
- [25] M Tromp, Benno Lansdorp, CP Bleeker-Rovers, JM Klein Gunnewiek, BJ Kullberg, and P Pickkers. "Serial and panel analyses of biomarkers do not improve the prediction of bacteremia compared to one procalcitonin measurement". *Journal of infection* 65.4 (2012), pp. 292–301.
- [26] Seri Jeong, Yongjung Park, Yonggeun Cho, and Hyon-Suk Kim. "Diagnostic utilities of procalcitonin and C-reactive protein for the prediction of bacteremia determined by blood culture". *Clinica Chimica Acta* 413.21-22 (2012), pp. 1731–1736.

- [27] Joseph Guillén, Jiankun Liu, Margaret Furr, Tianyao Wang, Stephen Strong, Christopher C Moore, et al. "Predictive models for severe sepsis in adult ICU patients". *Systems and Information Engineering Design Symposium (SIEDS)*, 2015. IEEE. 2015, pp. 182–187.
- [28] Haibo He and Eduardo A Garcia. "Learning from imbalanced data". *IEEE Transactions on knowledge and data engineering* 21.9 (2009), pp. 1263–1284.
- [29] Tom Fawcett. "An introduction to ROC analysis". *Pattern recognition letters* 27.8 (2006), pp. 861–874.
- [30] Jesse Davis and Mark Goadrich. "The relationship between Precision-Recall and ROC curves". *Proceedings of the 23rd international conference on Machine learning*. ACM. 2006, pp. 233–240.
- [31] Rajul Parikh, Annie Mathai, Shefali Parikh, G Chandra Sekhar, and Ravi Thomas. "Understanding and using sensitivity, specificity and predictive values". *Indian journal of ophthalmology* 56.1 (2008), p. 45.
- [32] Alexandru Niculescu-Mizil and Rich Caruana. "Predicting good probabilities with supervised learning". *Proceedings of the 22nd international conference on Machine learning*. ACM. 2005, pp. 625–632.
- [33] Franz Ratzinger, Michel Dedeyan, Matthias Rammerstorfer, Thomas Perkmann, Heinz Burgmann, Athanasios Makristathis, et al. "Neither Single nor a Combination of Routine Laboratory Parameters can Discriminate between Gram-positive and Gram-negative Bacteremia". *Scientific reports* 5 (2015), p. 16008.
- [34] Leonard Leibovici, Mical Paul, Anders D Nielsen, Evelina Tacconelli, and Steen Andreassen. "The TREAT project: decision support and prediction using causal probabilistic networks". *International journal of antimicrobial agents* 30 (2007), pp. 93–102.
- [35] Subramani Mani, Asli Ozdas, Constantin Aliferis, Huseyin Atakan Varol, Qingxia Chen, Randy Carnevale, et al. "Medical decision support using machine learning for early detection of late-onset neonatal sepsis". *Journal of the American Medical Informatics Association* 21.2 (2014), pp. 326–336.
- [36] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, et al. "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.
- [37] Bernard Hernandez, Pau Herrero, Timothy M. Rawson, Luke S. P. Moore, Esmita Charani, Alison H. Holmes, et al. "Data-driven Web-based Intelligent Decision Support System for Infection Management at Point-Of-Care: Case-Based Reasoning Benefits and Limitations". 5 (2017), pp. 119–127.
- [38] Alexander R Levine, Midori Tran, Jonathan Shepherd, and Edgar Naut. "Utility of initial procalcitonin values to predict urinary tract infection". *The American journal of emergency medicine* (2018).
- [39] Rui-Ying Xu, Hua-Wei Liu, Ji-Ling Liu, and Jun-Hua Dong. "Procalcitonin and C-reactive protein in urinary tract infection diagnosis". *BMC urology* 14.1 (2014), p. 45.

- [40] Philipp Schuetz, Werner Albrich, and Beat Mueller. "Procalcitonin for diagnosis of infection and guide to antibiotic decisions: past, present and future". *BMC medicine* 9.1 (2011), p. 107.
- [41] Erica M Caffarini, Joshua DeMott, Gourang Patel, and Ishaq Lat. "Determining the clinical utility of an absolute procalcitonin value for predicting a positive culture result". *Antimicrobial agents and chemotherapy* (2017), AAC-02007.
- [42] Nicolas Moser, Jesus Rodriguez-Manzano, Tor Sverre Lande, and Pantelis Georgiou. "A scalable ISFET sensing and memory array with sensor auto-calibration for on-chip real-time DNA detection". *IEEE transactions on biomedical circuits and systems* 12.2 (2018), pp. 390–401.
- [43] Jesus Rodriguez-Manzano, Ahmad Moniri, Kenny Malpartida-Cardenas, Jyothsna Dronavalli, Frances Davies, Alison Holmes, et al. "Simultaneous single-channel multiplex and quantification of carbapenem-resistant genes using multidimensional standard curves". *bioRxiv* (2018), p. 409912.

CONCLUSIONS AND FUTURE PERSPECTIVE

This chapter discusses the benefits of introducing an intelligent CDSS in health care: the convenience of combining various sources of data at the point of care (subsections 7.1.1 and 7.1.2), the importance of local AMR surveillance (subsection 7.1.3) and the need to adapt CDSSs to the pathway followed by clinicians in infection management (subsection 7.1.4). After this, the chapter summarises the overall contribution and describes the individual contribution of each decision support module to provide enhanced care (section 7.2). To conclude, it suggests potential research directions (section 7.3).

7.1 SATISFYING THE NEEDS IN INFECTION MANAGEMENT

Clinical decision support systems represent a vital component of the healthcare enterprise and have evolved from paper-based content-driven to digital systems that offer all medical stakeholders access to analytic tools for critical decision making. As such, Enhanced, Personalized and Integrated Care for Infection Management at Point of Care (EPiC IMPOC) has been designed to record a complete set of vital signs at the patient's bedside on hand-held computing devices while providing instant bed-side decision-making assistance to health care workers. The needs within infection management which have been satisfied through the implementation of EPiC IMPOC are discussed below.

7.1.1 *Need of integration with electronic health records*

For health care providers, electronic health records (EHR) systems are the preferred method of data implementation and has quickly become the backbone of healthcare organizations in high resource settings. As such, integration into EHR is no longer a luxury but a desired addition. This integration facilitates the collection of data to provide effective and efficient health care. As such, EPiC IMPOC feeds from three databases within the National Health Services (NHS). Firstly, the patient administration system provides demographics such as age or gender and hospital management information such as insurance number, date of admission or admitted ward. The pathology laboratory system contains biochemical markers results such as bilirubin, c-reactive protein or white cell counts. Last, the microbiology system provides susceptibility test outcomes to identify the presence of pathogens and determine which antimicrobials are effective. Access to vital sign recordings, symptoms and organ support was not provided by the NHS during this research. Thus, such information has been collected through EPiC IMPOC using

a standardised form (see C3 in Table 3.2). The variables were defined by clinicians and infection specialists.

7.1.2 Need of a point-of-care decision support system

The aforementioned adoption of EHR promotes the development of point of care (POC) technologies which are meant to minimize the time spent on documentation and provide decision support that fits in the workflow. Furthermore, a wealth of research has demonstrated the importance of effective communication to promote continuity and consistency of care in health centres [1]. This is crucial to boost the quality of the health care perceived by patients which is highly dependent on establishing trusting, empathetic and reliable relationships [2]. As such, EPiC IMPOC has been designed to be accessible at the point of care, facilitate surveillance, ensure rapid disease diagnosis and improve patient interaction, management and monitoring [3]. Moreover, it stimulates discussion and promotes shared decision making.

7.1.3 Need of understanding local AMR patterns

The majority of clinical decision support systems and other independent research addressing antimicrobial therapy selection often rely on national or even worldwide guidelines and policies. As such, these often neglect the variability caused by external factors such as hospital settings or geographical location. The prevalence of infectious pathogens vary geographically with disparate levels of resistance rates around the world [4]. Hence, it is necessary to understand local antimicrobial resistance patterns to prescribe appropriate empirical therapies. On the other hand, the resources available such as antimicrobials, microbiology laboratory equipment or the susceptibility test guidelines used might also differ among health care centres. For this reason, EPiC IMPOC includes a module to automatically compute from susceptibility test data AMR statistics such as the prevalence of pathogens (see graphs in section 4.3), the resistance rates and trends (see tables in section 4.3) or the antimicrobial spectrum of activity (see section 4.5) among others.

7.1.4 Need to align with the infection management pathway

The acute infection management pathway followed by clinicians (see Figure 7.1) has been defined as a stepwise Bayesian model in which each step adds systematically information to optimise diagnosis and management of infection [5]. In summary, clinicians begin with a predefined risk of infection being present and add further information to optimise decisions on diagnosis and management of the patient. First, they evaluate the physiological parameters of the patient and try to localise and confirm the infection. Then, physicians review and plan further investigations such as pathology laboratory tests, specimen collection for susceptibility testing or imaging. These previous steps allow physicians to construct a clinical picture of the severity of the infection. From this clinical picture, a decision to whether or not to initiate antimicrobial therapy is made with local microbiology guidance as a determinant factor. Finally, an internal and/or external review of the patient is conducted to refine the antimicrobial therapy and the entire process is repeated.

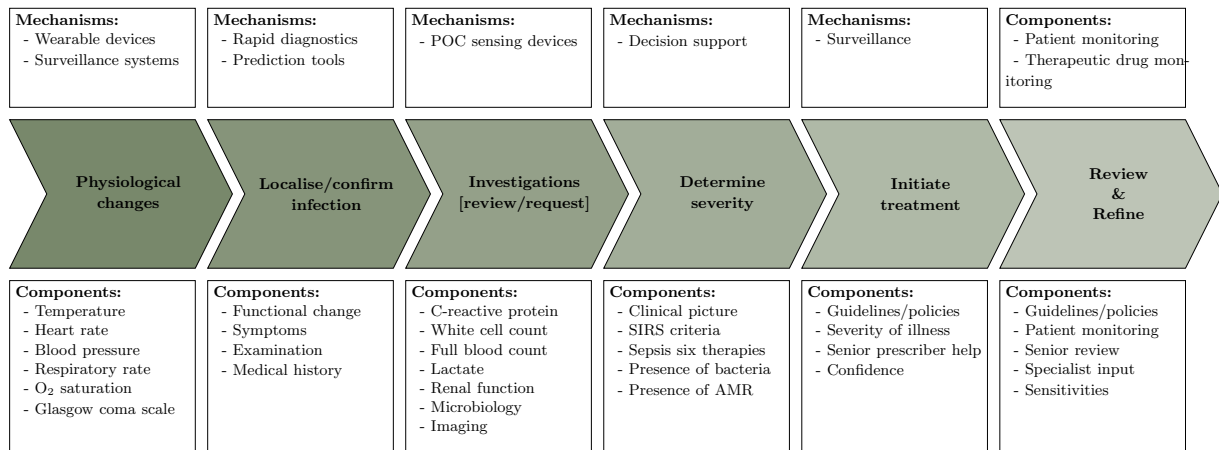


Figure 7.1: Acute infection management pathway. The stepwise Bayesian like infection management pathway followed by clinicians as described in [5]. Boxes below describe parameters commonly required during such phases of the infection management pathway. Boxes above suggest methods, systems or techniques to improve data collection on such phases. The module(s) providing support on each of the steps are: case-based reasoning (CBR), probabilistic inference (PI) and antimicrobial resistance surveillance (AMR).

How does EPiC IMPOC fit in the infection management pathway?

In the first phase of the infection management pathway, the physician reviews the physiological changes and the medical history of the patient. For such purpose, previous cases of the patient are shown within the CBR module (see D3 in Table 3.3). Moreover, the physiological parameters and other detected symptoms are recorded using the data collection form (see C3 in Table 3.2). After this, physicians try to confirm and localise the infection to review and plan further investigations. Consequently, the probabilistic inference module provides the likelihood of positive culture and the most plausible sites (see D6 in Table 3.2). This information provides a prior estimate of the presence of infection and assists physicians to plan further investigations such as specimen collection for susceptibility testing. This is particularly useful if the patient does not have localised signs or symptoms of infection but infection is suspected. In order to determine the severity of the infection and aid antimicrobial therapy selection, the presence of bacteria (see C5 in Table 3.2) and antimicrobial resistance (see resistance rates in section 4.3) are examined. From this clinical picture, a decision to whether or not to initiate antimicrobial therapy is made based on guidelines, previous experience and further senior prescriber help. This knowledge base is provided to physicians as a set of past similar cases, applied treatments and corresponding outcomes (see D4 in Table 3.3). To conclude, the patient is monitored and reviewed (see C4 in Table 3.2) to refine the antimicrobial therapy.

7.2 OVERALL DISCUSSION AND CONTRIBUTIONS

The main contribution within this research is EPiC IMPOC, an intelligent decision support system for infection management at the point of care. As discussed in the literature review, the majority of decision support systems focus on a single domain (e.g. bloodstream infection) and a single type of decision support (e.g. antimicrobial therapy selection). In contrast to these, EPiC IMPOC assists clinicians on multiple areas of the infection management pathway. The specific contributions within EPiC IMPOC (see Figure 7.2) are discussed below. Note that the literature review has been previously summarised (see section 2.3) and the full content is provided as supplementary material (see appendices B.1 and B.2).

■ *A hybrid clinical decision support system*

This thesis presents a novel clinical decision support system for infection management to provide personalized, accurate and effective diagnosis at point of care. The proposed system, which has been denoted as EPiC IMPOC, incorporates two main decision support engines: case-based reasoning to facilitate vital sign collection, patient monitoring and further inspection of past similar cases and probabilistic inference to provide stepwise guidance within the infection management pathway followed by clinicians. Overall, the system facilitates access and visualization of patient data (e.g. pathology/microbiology laboratory results) while providing tailored decision support (e.g. diagnosis or therapy selection). EPiC IMPOC has been implemented as a web-based platform that can be accessed at the point of care from computers or hand-held devices. The system usability score was 68.5 which is above average with margin for improvement.

■ *Promoting patient engagement in infection management*

An antimicrobial information module co-designed with patients has been integrated in EPiC IMPOC to provide personalised antimicrobial information in the form of a leaflet [6]. The provision of this information has demonstrated better short-term understanding of patients on infections and antimicrobial use [7].

■ *What is the likelihood of a positive culture?*

Previous research has focused predominantly on bacteremia; that is, bloodstream infections. To the best of the author's knowledge, only one study has addressed overall culture positivity regardless of the site of infection [8]. The study applied the Youden's index on the procalcitonin levels of 519 patients and reported area under the ROC was 0.62. The sensitivity and specificity values were not specified. In this research, the best performance was obtained using an artificial neural network to infer culture positivity from six biochemical markers. The area under the ROC was 0.91 with sensitivity and specificity scores of 0.78 and 0.88 respectively.

Table 7.1: Positive culture: existing research.

	n	AUCROC	SENS	SPEC
Positive [8]	1	0.62	-	-

Table 7.2: Positive culture: this research.

	n	AUCROC	SENS	SPEC
Positive [9]	6	0.91	0.78	0.88

■ *What is the most plausible site of the infection?*

The majority of studies within the existing literature focus on a particular type of infection. In bloodstream infections, the reported area under the ROC was 0.75 (± 0.08) using regression analysis, 0.88 (± 0.08) using tree-based methods and 0.81 (± 0.09) using support vector machines. In general, the results reported within the existing literature present a significant imbalance between sensitivity and specificity. In this research, an area under the ROC for bacteremia detection of 0.96 was achieved with a reasonable sensitivity-specificity balance (0.86–0.93). The rest of culture sites are studied in a much lesser extent within the literature. For comparison purposes, the best results reported within the literature have been summarised (see Table 7.3). In this research, the same methodology was applied to the different sites of infection. The results obtained are quite consistent (see Table 7.4) and provide an appropriate trade off between the number of parameters required and the performance.

Table 7.3: Site of infection: existing research.

		n	AUCROC	SENS	SPEC
Blood	[10]	30	0.92	0.90	0.85
Urine	[11]	211	0.90	0.62	0.95
Wound	[12]	29	0.89	0.67	0.95
Sputum	[13]	1	-	0.73	0.65

Table 7.4: Site of infection: this research.

		n	AUCROC	SENS	SPEC
Blood	[14]	6	0.96	0.87	0.93
Urine	[14]	6	0.82	0.66	0.83
Wound	[14]	6	0.92	0.79	0.80
Sputum	[14]	6	0.93	0.83	0.87

■ *Is the infectious pathogen Gram-positive or Gram-negative?*

Overall, the prediction of the Gram status of a bacteria is not achievable with routinely collected biochemical markers. In bloodstream infections, the best study reports an area under the ROC of 0.68 [15]. However, the values of sensitivity (0.45) and specificity (0.80) are clearly imbalanced. The study presents a comparison of algorithms and parameters from which the K-Star method with 7 biochemical markers as input features was selected. The performance obtained in this research using six biochemical markers was very similar with an area under the ROC of 0.67. Furthermore, the values of sensitivity (0.60) and specificity (0.62) are balanced. Moreover, the conducted research has also evaluated the performance on other sites of infection. No similar studies were found in the existing literature for further comparison. From the results reported in this thesis it is possible to conclude that prediction of the Gram status of bacteria is not achievable with the selected routinely collected biochemical markers [15].

Table 7.5: Gram status: existing research.

		n	AUCROC	SENS	SPEC
Positive	n/a	-	-	-	-
Blood	[15]	7	0.68	0.45	0.80
Urine	n/a	-	-	-	-
Sputum	n/a	-	-	-	-
Wound	n/a	-	-	-	-

n/a=no similar studies were found.

Table 7.6: Gram status: this research.

		n	AUCROC	SENS	SPEC
Positive	†	6	0.56	0.51	0.57
Blood	†	6	0.67	0.60	0.62
Urine	†	6	0.62	0.53	0.62
Sputum	†	6	0.66	0.58	0.66
Wound	†	6	0.65	0.67	0.51

†=results presented in section 6.7.

- *What are the most plausible genera/species?*

The prediction of the genus or species of an infectious pathogen using routinely collected biochemical markers has not been addressed in this research. Note that this problem is more challenging than the identification of the Gram status of a bacteria which was not accurate with the selected biomarkers. As such, to provide physicians with an insight of the infectious pathogens that might be causing the infection before the results of susceptibility tests are available, various site specific AMR statistics have been computed (see section 4.3) such as the prevalence of pathogens.

- *What antimicrobials are effective to treat the infection?*

The case-based reasoning methodology has been included to inform physicians of previous past cases to assist in antimicrobial therapy selection. In order to evaluate the validity and usability of the system, a pilot study was conducted in the ICU. The system advocated the same results as those suggested by infection specialists in 84% of the cases. Furthermore, participants highlighted the utility of the case-based reasoning engine to promote knowledge transfer among health care professionals and the benefits of having access to real-time patient data at the point of care.

In addition, a number of AMR-statistics have been computed to promote awareness among clinicians and potentially assist them in antimicrobial therapy selection. Firstly, the local antimicrobial resistance rates are used to assess the effectiveness of the antimicrobials (see section 4.4). These resistance rates have been combined to compute the antimicrobial spectrum of activity index (see section 4.5). The proposed index overcomes the ambiguity of the commonly used narrow-broad notation and better assist clinicians to select empirical antimicrobial therapies. Finally, the antimicrobial resistance trend has been computed to inform of local resistance patterns (see section 4.6).

- *Antimicrobial resistance rates*

The local resistance rates for specific pathogen-antimicrobial pairs have been computed for each sample type (e.g. urine or blood). The graphical representation provided clusters antimicrobials and pathogens based on the similarity of the resistance profiles. Thus, pathogens and antimicrobials with similar resistance behaviour are placed together. From these clusters, a number of patterns within susceptibility testing arose which could be exploited to improve prescription practices.

- *Antimicrobial resistance trends*

The local resistance trends for specific pathogen-antimicrobial pairs have been computed through regression analysis. As such, the single antimicrobial resistance trend has been defined to provide a continuous numerical value within the range $[-1,1]$. Thus, the secular resistance trends are quantified numerically and it is possible to assess their significance by inspecting the confidence interval. The results obtained for three different scenarios demonstrated that the obtained trends align with those reported in the literature.

- *Enhanced antimicrobial spectrum of activity*

Empirical therapies commonly rely on broad spectrum antimicrobials to provide treatment active to most likely microbes. As such, these therapies need to cover all potential resistant pathogens. The existing categorization based on the terms broad and narrow applied to the spectrum of antimicrobials has been widely misused and

needs revision. To resolve such ambiguity, the antimicrobial spectrum of activity index has been defined to provide a continuous numerical value within the range $[0,1]$. The results obtained align with those reported in the literature and provide deeper understanding into the antimicrobial coverage for each of the Gram statuses.

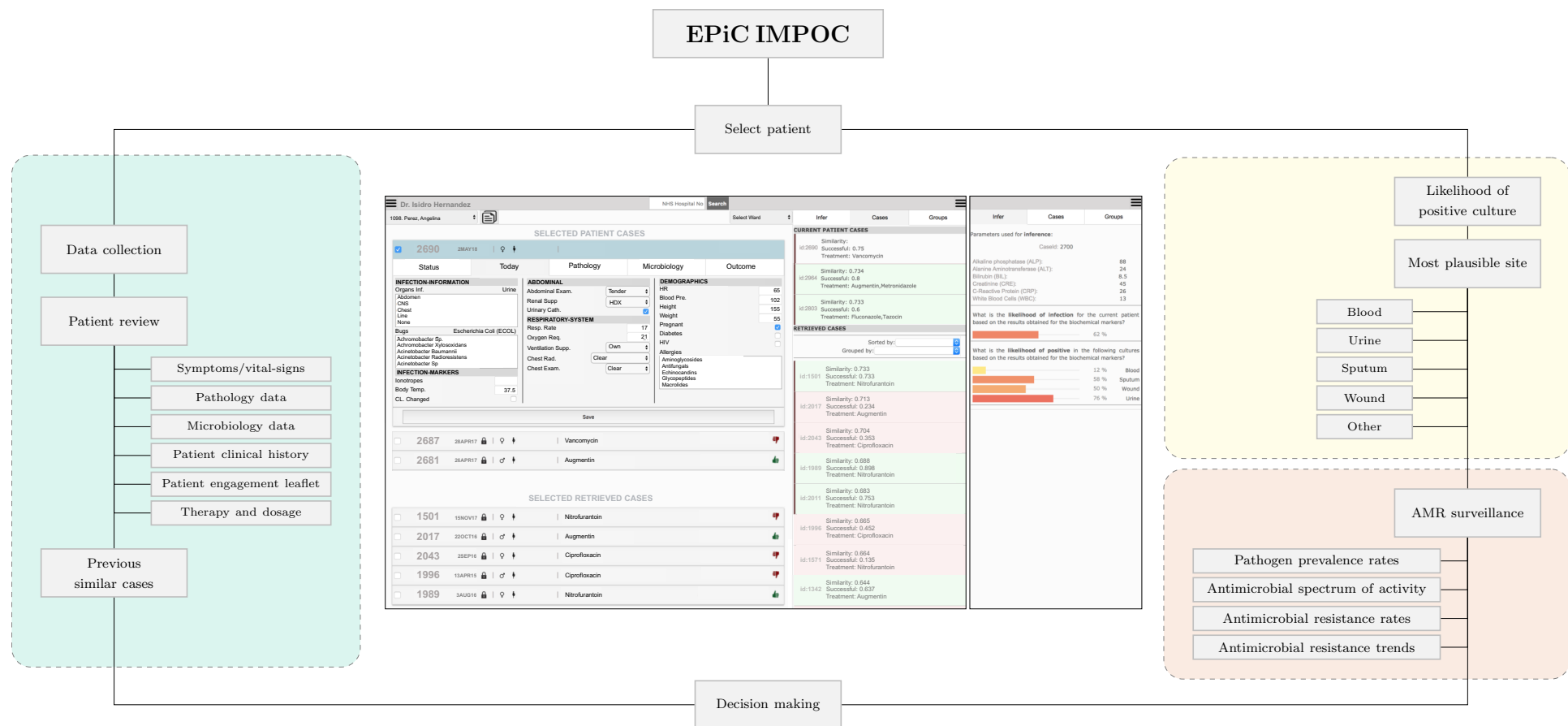


Figure 7.2: EPiC IMPOC graphical summary. Diagram summarizing the key aspects and contributions of EPiC IMPOC. The diagram is composed of four main sections: the user interface (see image in the middle), the case-based reasoning module (see green background), the probabilistic inference module (see yellow background) and the AMR surveillance module (see orange background). The case-based reasoning module facilitates data collection, patient review and further inspection of past similar cases. The probabilistic inference module provides step wise decision support to assist clinicians estimate the likelihood of infection and the most plausible sites to plan further investigations. The AMR surveillance module provides a number of statistics such as pathogen prevalence, antimicrobial spectrum of activity, antimicrobial resistance rates and antimicrobial resistance trends to further assist in antimicrobial therapy selection. All this information is provided to clinicians to enhance decision making at the point of care.

7.3 FUTURE RESEARCH DIRECTIONS

The future research recommendations presented in this section are organized according to the decision support modules implemented in EPiC IMPOC (see Chapter 3). The main objectives of these modules are: (i) the effective communication of local AMR patterns (ii) the facilitation of patient history and previous similar cases for further inspection and (iii) the provision of targeted decision support within the infection management pathway.

7.3.1 *Effective communication of local AMR patterns*

EPiC IMPOC includes a decision support module to compute AMR patterns from susceptibility test data (see Chapter 4). At the moment, it presents the proportion of type of cultures requested, the resistance rates and trends for microbe-antimicrobial pairs and the overall antimicrobial spectrum of activity. Moreover, it displays the temporal evolution and provides forecast insights. This information should be presented in a comprehensive manner to health care workers. As such, developing an effective communication strategy that meets the needs of the broad range of stakeholders (e.g. nurses, clinicians, infection specialists, public health agencies or the public) is clearly an important research area. A potential approach relies on the design of a dynamic and interactive data visualization tool (e.g. dashboard) within the web browser¹.

7.3.2 *Improving history review and case retrieval*

EPiC IMPOC includes a decision support module to facilitate patient history review and inspection of previous similar cases (see Chapter 5). In overall, the development of this module has been very limited due to a lack of clinical data.

Addition of parameters within the patient case

The clinical record of a patient defined in the ICU trial contains a wide variety of clinical features such as demographics, symptoms, vital signs, laboratory results and other investigations (see Appendix E.1). The selection of the features was carried out by infection specialists. The clinical variables selected were consistent with the views and perceptions of other infection specialists and health care staff. However, from the feedback provided after the trial further parameters could have been potentially included (see section 5.5.2). Note that the relevant clinical parameters vary slightly among cohorts.

Defining the case similarity

The case-based reasoning (CBR) methodology presented in this thesis uses standards distance metrics to retrieve the most similar cases within the case base. However, there are other techniques to define the similarity between two cases that could be investigated (e.g. fuzzy logic [16]). Also, the feature weights used in the presented clinical trials was selected by the infection specialists. The reason for the inability to fine tune the features was the lack of clinical data.

¹<https://d3js.org>

7.3.3 Enhancing stepwise decision support in infection management

EPiC IMPOC includes a decision support module to provide timely and targeted decision assistance for infection management (see Chapter 6). As such, this module aims to guide clinicians through the different steps of the infection management pathway (see Figure 7.1). At the moment, it computes the likelihood of infection and the most plausible sites of infection. For this purpose, it uses six clinically relevant biochemical markers which are routinely collected. Future research should focus on (i) analyse the individual contribution of each of the biochemical makers to answer the previously postulated questions (ii) explore and include biochemical markers that are specific for infectious diseases (iii) determine the combination of biochemical markers which provides the best performance for each condition. Note that for practicability reasons only a reduced set of parameters which are commonly available have been considered. Yet, the selection of parameters entirely depends on the health care centre in which the system is deployed.

In clinical domains, the vital signs and laboratory tests collected often vary depending on the severity of the disease or the resources available. Note that these laboratory tests carry certain costs to the health care centres. As such, it would be extremely useful to determine whether missing parameters can be estimated accurately from those that are available. Firstly, this would provide additional information on the patient condition to clinicians. More importantly, this would help addressing one of the most common problems of innovation adoption in health care: the usability and performance of the algorithms in scenarios with missing data. At the moment, missing data is imputed using the median. However, imputing missing parameters more accurately has the potential to boost the performance of algorithms and promote their use in clinical practice.

Table 7.7: Description of potential scores to include in EPiC IMPOC.

SIRS	Sepsis	Severe sepsis
The systemic inflammatory response syndrome (SIRS) is an inflammatory state affecting the whole body. The causes are broadly classified as infectious and non infectious. This condition is often related with organ dysfunction or failure and sepsis. However, many experts consider SIRS to be overly sensitive [17]. Parameters: – body temperature (T) – heart rate (hr) – respiratory rate (rr) – white blood cell counts (wbc) Rule: at least two of the following – $T > 38^{\circ}\text{C}$ or $T < 36^{\circ}\text{C}$ – $\text{hr} > 90$ – $\text{rr} > 20$ or $\text{PaCO}_2 < 32\text{mmHg}$ – $\text{wbc} > 12000/\text{mm}^3$ or $\text{wbc} < 4000/\text{mm}^3$	Sepsis occurs when the chemicals released by the immune system in the bloodstream as a response to an infection cause life threatening organ dysfunction [18]. Parameters: – SIRS – suspected infection – confirmed infection Rule: – SIRS criteria with either suspected or confirmed infection.	Severe sepsis is a complication of sepsis in which this causes poor organ function or insufficient blood flow. It is associated with acute circulatory, cellular and metabolic abnormalities with greater risk of mortality than sepsis alone [18]. Parameters: – sepsis – organ damage – systolic blood pressure – lactate Rule: – sepsis criteria – signs of organ damage – hypo-tension ($\text{sbp} < 90$) – lactate $> 4\text{mmol}$

Keys: mmHg =millimeter of mercury; mm^3 =cubic millimeter; mmol =millimoles; PaCO_2 =partial pressure of carbon dioxide.

To conclude, further research could focus on adapting the stepwise decision support module to align with the more specific sepsis management pathway. A number of guidelines for the management of the sepsis pathway have recently been published [19]. These guidelines agree on the importance of early detection (0 to 6 hours), lung-protective ventilation [20] and goal-directed therapies [21]. The therapeutic plan to manage sepsis

consists on the following main steps: clinical evaluation, laboratory evaluation and management [21]. Firstly, vital signs and symptoms are used to assess the airway, breathing and circulation of the patient. In addition, the clinician identifies signs of the systemic inflammatory response syndrome (SIRS) and possible organ dysfunction. The assessment of renal function (creatinine) and hepatic function (bilirubin and alkaline phosphatase) are already included in EPiC IMPOC. However, other biochemical markers such as platelets, blood gas or lactate could be included to detect other conditions such as coagulation. The inclusion of this additional information will also enable to automatically compute common scores used for bloodstream infection (see Table 7.7). Because the site of infection and infectious pathogens are usually not known initially in a patient with sepsis, microbiology cultures are requested and often broad-spectrum antibiotics administered [22, 23]. Several observational studies indicate that outcomes of sepsis [24] and septic shock [22] are worse if the causative pathogens are not sensitive to the initial antibiotic regimen. Thus, supporting empirical and personalised antimicrobial therapy selection at early stages would be of great help. To conclude, identification of the infection source and further monitoring is required to narrow and refine the antimicrobial regimen.

7.3.4 *Integration of further technology*

Physiologic monitoring systems were introduced into the ICU in the 1970s and have not changed substantially since then [25]. The most common physiologic parameters displayed on ICU monitors include blood pressure, oxygen saturation, heart rate and respiratory rate. These physiologic parameters are monitored on a regular basis by nurses to assure stability. Timely detection of changes is a common challenge for nurses who have to integrate 10 or more rapidly changing parameters. Furthermore, in the case of an unexpected and potentially life-threatening event, the cognitive demand to identify these changes reduces available cognitive resources for other important tasks such as communication with other physicians or taking corrective actions. A potential solution relies on analysing these physiological parameters in real-time to detect drastic and potentially harming changes. For this purpose, all the physiologic monitoring should be automatically recorded within the health care records.

EPiC IMPOC is integrated with the electronic health records and can analyse and deliver the data to any device connected to the hospital network. Thus, further methodologies to analyse the physiological parameters could be integrated into the system to alert physicians of possible harming changes. Moreover, the system would be able to communicate to health care workers that certain laboratory tests (e.g. pathology or microbiology) have become available.

7.3.5 *Implementation in low- and middle-income countries*

The burden of AMR has received significant attention from high income countries yet the vast majority of deaths are projected to occur in low- and middle-income countries which generally have to face the lack of financial, material and human resources within their health care systems [4, 26]. In the last years, the introduction of decision support systems to improve health care provision in settings with limited resources has gained recognition [27–31]. One of the main limitations is the weak capacity of health systems and the shortage of qualified health workers that leads to the delivery of tasks to a lower

cadre such as auxiliary nurses and community health workers. While task shifting has been recognised for improving efficiency and access to health services, concerns exist to whether lower cadre health workers are competent and equipped to effectively handle those additional responsibilities [32]. As such, there is potential to implement the presented decision support system to enhance the delivery of evidence-based medicine. The main critical factor is the revision of cases for further inclusion into the case base. Note that retaining cases which were not optimally managed could lead to poor antimicrobial use. Ideally, the aim of the case base is to represent a gold standard. However, the definition of the gold standard evolves and the case base should evolve with it. Thus, it would be extremely useful a comprehensive assessment of the needs and outcomes of clinical decision support systems within this new user base to determine the best strategy to update the case base. This information would also inform implementers and policy makers on the effectiveness of the presented technology were is needed the most.

7.3.6 Summary of potential research directions

The potential research directions previously discussed have been summarised in this section (see Table 7.8). These ideas include both short-and long-term goals to enhance EPiC IMPOC and therefore maximise usability, performance and adherence. Note that these future pieces of work also include the vision of the author regarding the integration of existing and future diagnostic and monitoring technologies into one system. In addition to the points described in Table 7.8, future research should focus on evaluating each of the previously mentioned components within a number of health care centres and cohorts during long periods to identify potential areas of improvement.

Table 7.8: Summary of future research directions.

AMR surveillance	Case-based reasoning	Probabilistic inference
Develop a communication strategy to present in a comprehensive fashion all the computed AMR statistics. For example, a dynamic and interactive web-based visualization tool.	Evaluate various distance metrics and fine tune the feature weights . Introduce the concept of preference-based solution to leverage the information retrieval outcomes.	Assess the individual contribution of each of the selected biochemical makers in the different scenarios. Determine biomarkers that provide the best performance. Note that these may vary among scenarios.
Diagnostic technology	Antimicrobial dosing	
Integrate diagnostic tests that help identifying the infectious pathogen and its concentration. Integrate monitoring devices so that all the vital signs are collected and stored automatically into the electronic health records without human intervention.	Integrate PK/PD models that use information of the patient, the pathogen and the antimicrobial in conjunction with the dynamics of body to calculate the optimal antimicrobial regime including both dosage and time.	Identify variables which could be incorporated into the models to boost their performance. Estimate missing vital signs or biochemical markers from those parameters which are already available. Compute scores used by physicians to determine the condition of the patient and the severity of the infection.

Keys: AMR=antimicrobial resistance; PK/PD=pharmacokinetic/pharmacodynamic;

7.4 FINAL CONCLUSION

The research presented in this thesis describes the design and implementation of a clinical decision support system for infection management that includes a number of machine learning techniques. The system, which has been denoted as EPiC IMPOC, incorporates two main decision support engines: case-based reasoning to facilitate patient

monitoring and further inspection of past similar cases and probabilistic inference to provide stepwise guidance within the infection management pathway followed by clinicians. The participants highlighted the utility of the case-based reasoning engine to promote knowledge transfer among health care professionals and the benefits of having access to real-time patient data at the point of care. In addition, the system provides the likelihood of positive culture and the most plausible sites with a high degree of confidence as demonstrated by the results presented in this thesis. Moreover, a number of local AMR statistics are computed automatically from susceptibility test data to promote education and awareness among physicians. All these elements combined result in a state-of-the-art clinical decision support system which assists physicians on multiple areas within infection management to facilitate the provision of evidence-based and personalized medicine.

The research presented in this thesis sets the basis of a novel type of comprehensive decision support systems that analyse local hospital data to better assist clinicians through the infection management pathway. Therefore, EPiC IMPOC shows potential to revamp clinician's prescription practices to reduce considerably further development of AMR.

BIBLIOGRAPHY

- [1] Michael Leonard, Suzanne Graham, and Doug Bonacum. "The human factor: the critical importance of effective teamwork and communication in providing safe care". *BMJ Quality & Safety* 13.suppl 1 (2004), pp. i85–i90.
- [2] Anthony C Berman and Darryl S Chutka. "Assessing effective physician-patient communication skills: "Are you listening to me, doc?"". *Korean journal of medical education* 28.2 (2016), p. 243.
- [3] Bernard Hernandez, Pau Herrero, Timothy M. Rawson, Luke S. P. Moore, Esmita Charani, Alison H. Holmes, et al. "Data-driven Web-based Intelligent Decision Support System for Infection Management at Point-Of-Care: Case-Based Reasoning Benefits and Limitations". 5 (2017), pp. 119–127.
- [4] Jim O'Neill. "Antimicrobial resistance: tackling a crisis for the health and wealth of nations". *Review on antimicrobial resistance* (2014), pp. 1–16.
- [5] Timothy Miles Rawson, Esmita Charani, Luke Stephen Prockter Moore, Bernard Hernandez, Enrique Castro-Sánchez, Pau Herrero, et al. "Mapping the decision pathways of acute infection management in secondary care among UK medical physicians: a qualitative study". *BMC medicine* 14.1 (2016), p. 208.
- [6] Timothy M Rawson, Luke SP Moore, Bernard Hernandez, Enrique Castro-Sanchez, Esmita Charani, Pantelis Georgiou, et al. "Patient engagement with infection management in secondary care: a qualitative investigation of current experiences". *BMJ open* 6.10 (2016), e011040.
- [7] Timothy M Rawson, Luke SP Moore, Enrique Castro-Sanchez, Esmita Charani, Bernard Hernandez, Vivian Alivizda, et al. "Development of a patient-centred intervention to improve knowledge and understanding of antibiotic therapy in secondary care". *Antimicrobial Resistance & Infection Control* 7.1 (2018), p. 43.
- [8] Erica M Caffarini, Joshua DeMott, Gourang Patel, and Ishaq Lat. "Determining the clinical utility of an absolute procalcitonin value for predicting a positive culture result". *Antimicrobial agents and chemotherapy* (2017), AAC–02007.
- [9] Bernard Hernandez, Pau Herrero, Timothy Miles Rawson, Luke SP Moore, Benjamin Evans, Christofer Toumazou, et al. "Supervised learning for infection risk inference using pathology data". *BMC medical informatics and decision making* 17.1 (2017), p. 168.
- [10] Qingqing Mao, Melissa Jay, Jana L Hoffman, Jacob Calvert, Christopher Barton, David Shimabukuro, et al. "Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU". *BMJ open* 8.1 (2018), e017833.

-
- [11] R Andrew Taylor, Christopher L Moore, Kei-Hoi Cheung, and Cynthia Brandt. "Predicting urinary tract infections in the emergency department with machine learning". *PloS one* 13.3 (2018), e0194085.
- [12] Pao-Jen Kuo, Shao-Chun Wu, Peng-Chen Chien, Shu-Shya Chang, Cheng-Shyuan Rau, Hsueh-Ling Tai, et al. "Artificial neural network approach to predict surgical site infection after free-flap reconstruction in patients receiving surgery for head and neck cancer". *Oncotarget* 9.17 (2018), p. 13768.
- [13] Anette Holm, Joergen Nexoe, Lene A Bistrup, Svend S Pedersen, Niels Obel, Lars P Nielsen, et al. "Aetiology and prediction of pneumonia in lower respiratory tract infection in primary care". *Br J Gen Pract* 57.540 (2007), pp. 547–554.
- [14] Beranrd Hernandez. "Pending" (2018).
- [15] Franz Ratzinger, Michel Dedeyan, Matthias Rammerstorfer, Thomas Perkmann, Heinz Burgmann, Athanasios Makristathis, et al. "Neither Single nor a Combination of Routine Laboratory Parameters can Discriminate between Gram-positive and Gram-negative Bacteremia". *Scientific reports* 5 (2015), p. 16008.
- [16] Didier Dubois, Henri Prade, Francesc Esteva, Pere Garcia, Lluís Godo, and Ramon López de Màntaras. "Fuzzy set modelling in case-based reasoning". *International Journal of Intelligent Systems* 13.4 (1998), pp. 345–373.
- [17] Janet M Lord, Mark J Midwinter, Yen-Fu Chen, Antonio Belli, Karim Brohi, Elizabeth J Kovacs, et al. "The systemic immune response to trauma: an overview of pathophysiology and treatment". *The Lancet* 384.9952 (2014), pp. 1455–1465.
- [18] Mervyn Singer, Clifford S Deutschman, Christopher Warren Seymour, Manu Shankar Hari, Djillali Annane, Michael Bauer, et al. "The third international consensus definitions for sepsis and septic shock (Sepsis-3)". *Jama* 315.8 (2016), pp. 801–810.
- [19] R Phillip Dellinger, Mitchell M Levy, Jean M Carlet, Julian Bion, Margaret M Parker, Roman Jaeschke, et al. "Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock: 2008". *Intensive care medicine* 34.1 (2008), pp. 17–60.
- [20] Acute Respiratory Distress Syndrome Network. "Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome". *New England Journal of Medicine* 342.18 (2000), pp. 1301–1308.
- [21] Emanuel Rivers, Bryant Nguyen, Suzanne Havstad, Julie Ressler, Alexandria Muzzin, Bernhard Knoblich, et al. "Early goal-directed therapy in the treatment of severe sepsis and septic shock". *New England Journal of Medicine* 345.19 (2001), pp. 1368–1377.
- [22] Emad H Ibrahim, Glenda Sherman, Suzanne Ward, Victoria J Fraser, and Marin H Kollef. "The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting". *Chest* 118.1 (2000), pp. 146–155.
- [23] L Leibovici, I Shraga, M Drucker, H Konigsberger, Z Samra, SD Pitlik, et al. "The benefit of appropriate empirical antibiotic treatment in patients with bloodstream infection". *Journal of internal medicine* 244.5 (1998), pp. 379–386.

- [24] Stephan Harbarth, Jorge Garbino, Jérôme Pugin, Jacques A Romand, Daniel Lew, and Didier Pittet. "Inappropriate initial antimicrobial therapy and its effect on survival in a clinical trial of immunomodulating therapy for severe sepsis". *The American journal of medicine* 115.7 (2003), pp. 529–535.
- [25] Frank A Drews. "Patient monitors in critical care: Lessons for improvement" (2008).
- [26] World Health Organization. "Global action plan on antimicrobial resistance". *Journal of global health* (2015).
- [27] Charles S Hall, Edward Fottrell, Sophia Wilkinson, and Peter Byass. "Assessing the impact of mHealth interventions in low-and middle-income countries—what has been shown to work?" *Global health action* 7.1 (2014), p. 25606.
- [28] Ibukun-Oluwa Omolade Adepoju, Bregje Joanna Antonia Albersen, Vincent De Brouwere, Jos van Roosmalen, and Marjolein Zweekhorst. "mHealth for clinical decision-making in sub-Saharan Africa: a scoping review". *JMIR mHealth and uHealth* 5.3 (2017).
- [29] Tilly A Gurman, Sara E Rubin, and Amira A Roess. "Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature". *Journal of health communication* 17.sup1 (2012), pp. 82–104.
- [30] Clara B Aranda-Jan, Neo Mohutsiwa-Dibe, and Svetla Loukanova. "Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa". *BMC public health* 14.1 (2014), p. 188.
- [31] Karin Källander, James K Tibenderana, Onome J Akpogheneta, Daniel L Strachan, Zelee Hill, Augustinus HA ten Asbroek, et al. "Mobile health (mHealth) approaches and lessons for increased performance and retention of community health workers in low-and middle-income countries: a review". *Journal of medical Internet research* 15.1 (2013).
- [32] Brent D Fulton, Richard MScheffler, Susan P Sparkes, Erica Yoonkyung Auh, Marko Vujicic, and Agnes Soucat. "Health workforce skill mix and task shifting in low income countries: a review of recent evidence". *Human resources for health* 9.1 (2011), p. 1.



LIST OF PUBLICATIONS

A.1 JOURNAL ARTICLES

———— Chapter 2 ————

Rawson TM, Moore LSP, **Hernandez B**, Charani E, Castro-Sanchez E, Herrero P, Hayhoe B, Georgiou P, and Holmes AH. A systematic review of clinical decision support systems for antimicrobial management: are we failing to investigate these interventions appropriately? *Clinical Microbiology and Infection*, 2017

Rawson TM, Charani E, Moore LSP, **Hernandez B**, Castro-Sanchez E, Herrero P, Georgiou P, and AH Holmes. Mapping the decision pathways of acute infection management in secondary care among uk medical physicians: a qualitative study. *BMC medicine*, 14(1):208, 2016

———— Chapter 3 ————

Rawson TM, Moore LSP, **Hernandez B**, Castro-Sanchez E, Charani E, Georgiou P, Ahmad R, and AH Holmes. Patient engagement with infection management in secondary care: a qualitative investigation of current experiences. *BMJ open*, 6(10):e011040, 2016

Rawson TM, Moore LSP, Castro-Sanchez E, Charani E, **Hernandez B**, Alividza V, Husson F, Toumazou C, Ahmad R, Georgiou P, and AH Holmes. Development of a patient-centred intervention to improve knowledge and understanding of antibiotic therapy in secondary care. *Antimicrobial Resistance and Infection Control*, 7(1):43, 2018

———— Chapter 4 ————

Hernandez B, Herrero P, Rawson TM, Moore LSP, Guemes A, Toumazou C, Holmes AH and Georgiou P. Resistance trend estimation using regression analysis to enhance antimicrobial surveillance: a multicentre study in London 2009-2016. *PLoS ONE* [submitted for publication], 2018

———— Chapter 5 ————

Hernandez B, Herrero P, Rawson TM, Moore LSP, Charani E, Holmes AH, and Georgiou P. Data-driven web-based intelligent decision support system for infection management at point-of-care: Case-based reasoning benefits and limitations. *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies*, 5:119–127, 2017

———— Chapter 6 ————

Hernandez B, Herrero P, Rawson TM, Moore LSP, Evans B, Toumazou C, Holmes AH, and Georgiou P. Supervised learning for infection risk inference using pathology data. *BMC medical informatics and decision making*, 17(1):168, 2017

A.2 CONFERENCE: ABSTRACTS AND POSTERS

Hernandez B, Herrero P, Rawson TM, Moore LSP, Charani E, Holmes AH and Georgiou P. Point of care intelligent decision support systems for antimicrobial prescribing in the intensive care unit. *EMBRACE conference on multidisciplinary approaches tackling AMR*, 2016

Rawson TM, Moore LSP, **Hernandez B**, Castro-Sanchez E, Charani E, Ahmad R, and Holmes AH. Patient and public engagement in antimicrobial stewardship: a stakeholder analysis of shared decision making during infection management. *Federation of Infection Society*, 2015

Moore LSP, Charani E, Herrero P, Georgiou P, **Hernandez B**, and Holmes AH. Case-based reasoning for antimicrobial prescribing decision support: A solution for critical care? *Medical Engineering Centres Annual Meeting and Bioengineering*, 2014

ICID

Hernandez B, Herrero P, Rawson TM, Moore LSP, Toumazou C, Holmes AH and Georgiou P. Enhancing antimicrobial surveillance: an automated, dynamic and interactive approach. *International Journal of Infectious Diseases (18th ICID)*, 73:122, 2018

Rawson TM, Moore LSP, **Hernandez B**, Castro-Sanchez E, Charani E, Ahmad R, and Holmes AH. Missed opportunities for shared decision making in antimicrobial stewardship: The potential consequences of a lack of patient engagement in secondary care. *International Journal of Infectious Diseases (17th ICID)*, 45:122–123, 2016

ECCMID

Rawson TM, **Hernandez B**, Blandy O, Moore L, Herrero P, Toumazou C, Sriskandan S, Georgiou P, and Holmes AH. Supervised machine learning for the prediction of bacteremia using routinely collected blood science data. *European Congress of Clinical Microbiology and Infectious Diseases (28th ECCMID)*, 2018

Rawson TM, **Hernandez B**, Blandy O, Moore L, Herrero P, Charani E, Gilchrist M, Toumazou C, Sriskandan S, Georgiou P, and Holmes AH. Case-based reasoning for individualized antimicrobial selection: can intelligent decision support improve antimicrobial management? *European Congress of Clinical Microbiology and Infectious Diseases (28th ECCMID)*, 2018

Rawson TM, Moore LSP, Castro-Sanchez E, Charani E, **Hernandez B**, Alividza V, Husson F, Toumazou C, Ahmad R, Georgiou P, and Holmes AH. Patient engagement with antimicrobial decision making in secondary care: a co-designed pilot intervention. *European Congress of Clinical Microbiology and Infectious Diseases (28th ECCMID)*, 2018

Rawson TM, **Hernandez B**, Moore LSP, Castro-Sanchez E, Charani E, Georgiou P, Ahmad R, and Holmes AH. Patient-centred interventions to promote citizen engagement with infection related decision making. *European Congress of Clinical Microbiology and Infectious Diseases (27th ECCMID)*, 2017

Rawson TM, Moore LSP, **Hernandez B**, Charani E, Castro-Sanchez E, Herrero P, Hayhoe B, Hope W, Georgiou P, and Holmes AH. Clinical decision support systems for antimicrobial management: a systematic review of interventions in primary and secondary care *European Congress of Clinical Microbiology and Infectious Diseases (27th ECCMID)*, 2017

Rawson TM, Charani E, Moore LSP, **Hernandez B**, Castro-Sanchez E, Herrero P, Georgiou P, and Holmes AH. Mapping decision pathways for acute infection management in UK secondary care: a qualitative study. *European Congress of Clinical Microbiology and Infectious Diseases (27th ECCMID)*, 2017

A.3 CONFERENCE: ORAL PRESENTATIONS

Hernandez B, Herrero P, Rawson TM, Moore LSP, Charani E, Holmes AH, and Georgiou P. Data-driven web-based intelligent decision support system for infection management at point-of-care: Case-based reasoning benefits and limitations. *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies*, 5:119–127, 2017

B

LITERATURE REVIEW

B.1 CLINICAL DECISION SUPPORT SYSTEMS FOR INFECTION MANAGEMENT

This section presents a literature review of the clinical decision support systems addressing infection management. For the sake of clarity, only those including the case-based reasoning methodology have been included. These systems stand in most cases as research standalone software with off-line evaluation replicating the scenarios encountered in clinical practice. The columns are explained in Table B.1. Moreover, the acronyms used within a table have been described in the keys section placed underneath.

Table B.1: Literature review: description of columns (infection diagnosis).

Name	Description
Reference	Link to the study for further information.
Year	Year in which the study was published.
Domain	The domain(s) addressed by the decision support system.
Purpose	The aim of the decision support system.
Method	The method or algorithm applied.
Number of variables (n)	The number of variables considered.
Type of variables	The type of variables considered. In those studies in which the variables were clearly defined, these were categorised into the following groups: symptoms, demographics, vital signs, biomarkers, patient history, microbiology and others. The number of variables was indicated between parenthesis.
Data	Description of the data.
Metrics	The metrics used to evaluate the system within the study.
Outcomes	The outcomes reported in the study.
Comments	Further comments and clarifications.

Ref	Year	Domain	Purpose	Platform	Method	n	Variables	Data	Metrics	Outcomes	Comments
[1]	2016	HCAI	classification	HER add-on	CBR (tree based) NLP (Naïve Bayes)	50	demographics, symptoms, biomarkers, microbiology, antibiotics prescribed, patient history, others + features extracted from clinicians narratives with NLP + decision tree provided by clinicians	2816 cases Bacteremia (149) Urinary (522) Enteric (39) Surgical (301) Respiratory (126) No infection (486) Extra-hospital (1084) Other (59) Cutaneous (50)	CBR overall accuracy kappa ACC KAPPA SENS SPEC ACC KAPPA SENS SPEC ACC KAPPA SENS SPEC ACC KAPPA SENS SPEC ACC KAPPA SENS SPEC ACC KAPPA SENS SPEC ACC KAPPA SENS SPEC ACC KAPPA SENS SPEC NLP accuracy	70.21% (70.19-70.22) 0.62 (0.605, 0.647) 97.80% 0.81 97.99% 97.79% 92.37% 0.78 96.74% 91.37% 99.33% 0.77 84.62% 99.53% 91.55% 0.66 93.36% 91.33% 97.12% 0.63 58.73% 98.92% 88.81% 0.57 57.61% 95.32% 77.84% 0.50 57.38% 90.65% 97.59% 0.43 45.76% 98.69% 98.01% 0.23 18.00% 99.46% modest performance	For the overall CBR accuracy and kappa coefficients the confidence intervals are also presented. The system stratifies patients according to the infection. Note that it has variables such as the doctor diagnosis or microbiology culture results (infectious pathogen and site), which provide a clear advantage (bias). These should not be used in the classification.
[2]	2010	ABM	diagnosis	research	CBR + rule based	81	binary symptoms (81)	30 simulated cases	accuracy	97%	
[3]	2012	Hepatitis	prediction of mortality	research	PSO CBRPSO KNN GNB SVM FDT	19	13 binary and 6 categorical (sex, steroid, antivirals, fatigue, liver big, liver firm, ascites, varices, ..., age, bil, alp, albumin, sgot, protime)	155 records (CP:32, CN:123)	PSO accuracy CBRPSO accuracy KNN accuracy GNB accuracy SVM accuracy FDT accuracy	82.66% 92.83% 83.45% 82.05% 86.92% 75.39%	The CBRPSO is the combination of CBR and PSO. The CBR method was used to assign weights to the attributes. The PSO method was use to perform binary clustering.
[4]	2013	UTI	diagnosis	research	rule-based (CBFCM)	30	24 binary and 6 categorical (age, fever, sex, pregnant, nausea, vomiting, diarrhea, hr, dysuria, ... + decision tree adjusted to clinical UTI guidelines)	174 positive UTI	CBFCM AGG ROC BNN AGG ROC	98% 0.88 (0.55-0.76) 86% 0.69 (0.55-0.76)	The AGG indicates the percentage of diagnose agreement with the prescription guidelines. The ROC includes the confidence intervals.
[5]	2010	HCAI	knowledge sharing	research	CBR		age, gender, symptoms, diagnosis, medical history and doctor identification	ARI (288) Gastroenteritis (40) Dermatitis (19) Rhinitis (16) UTI (11) Conjunctivitis (9) Others (220)	PREC REC	39.7% 68.09%	The study compares the medicines prescribed by the system with those prescribed using a user-focused evaluation method. The precision (PREC) and recall (REC) are presented.

keys: *HCAI*=healthcare associated infection; *ABM*=acute bacterial meningitis; *UTI*=urinary tract infection; *EHR*=electronic health records; *CBR*=case-based reasoning; *NLP*=natural language processing; *PSO*=particle swarm optimization; *KNN*=k-nearest neighbours; *GNB*=gaussian naive bayes; *SVM*=support vector machine; *FDT*=fuzzy decision theory; *CBFCM*=case-based fuzzy cognitive maps; *CP*=category positive; *CN*=category negative; *ARI*=acute respiratory infection; *ACC*=accuracy; *SENS*=sensitivity; *SPEC*=specificity; *PREC*=precision; *REC*=recall;

B.2 DIAGNOSIS OF INFECTIOUS DISEASES USING MACHINE LEARNING

This section presents a literature review of methods to diagnose infectious diseases from patient data. The reviewed studies stand as research articles without existing implementation in health care settings and have been divided into five groups according to the site of infection: bloodstream (see Tables B.4 and B.5), urinary tract (see Table B.6), respiratory tract (see Table B.7) and surgical site (see Table B.8). The columns are explained in Table B.3. Moreover, the acronyms used within a table have been described in the keys section placed underneath.

Table B.3: Literature review: description of columns (infection diagnosis).

Name	Description
Reference	Link to the study for further information.
Year	Year in which the study was published.
Inclusion criteria	Whether or not the study limited the inclusion of patients or episodes within the study. The most common restrictions were demographics (e.g. age>18), the fulfilment of established conditions (e.g. suspicion of infection by the physician or a certain score such as SIRS or SOFA). The majority of the studied also discarded those patients in which an antimicrobial regimen was applied.
Diagnosis criteria	Whether or not the study defined a clear criteria for the diagnosis of the infection. The most common used criteria was positive microbiology cultures.
Method	The method or algorithm applied.
Number of variables (n)	The number of variables considered.
Type of variables	The type of variables considered. In those studies in which the variables were clearly defined, these were categorised into the following groups: symptoms, demographics, vital signs, biomarkers, patient history, microbiology and others. The number of variables was indicated between parenthesis.
Data	Description of the data.
Outcomes	The outcomes presented in the study.
ROC	The area under the receiver operating characteristic curve.
SENS	The sensitivity or proportion of positives correctly classified as such.
SPEC	The specificity or proportion of negatives correctly classified as such.
Comments	Further comments and clarifications.

Table B.4: Literature review: diagnosis of bloodstream infection (I).

Ref	Year	Purpose	Inclusion c.	Sepsis c.	Method	n	Variables	Data	Outcomes	ROC	SENS	SPEC	Comments
[6]	2006	prediction	Y	N	CPN	214	symptoms, vital signs, demographics, biomarkers, microbiology	2514 patients derivation (790, 15.2% CP) evaluation (1724, 27.6% CP)	derivation validation	0.68 0.70	-- --	-- --	The study stratifies patients in three categories (low, medium and high) based on the risk of bacteraemia. The main objective of this categorization is to define a low risk group for whom blood cultures may not be needed.
[7]	2008	prediction	N	N	regression	10	Symptoms (4), vital signs (2), demographics (age), biomarkers (4)	2730 episodes derivation (2466, 8.3% CP) evaluation (1264, 8.0% CP)	derivation validation	0.80 0.75	0.98 0.97	0.29 0.29	
[8]	2012	prediction	Y	N	regression	1-2	biomarkers	342 patients (55, 16.08% CP)	PCT (0.253ug/L) IL6 (209pg/mL) LBP (2728ug/mL) CRP (148mg/L)	0.80 -- -- --	0.89 0.62 0.60 0.53	0.58 0.72 0.68 0.79	
[9]	2012	prediction	N	N	regression	4	PCT, CRP, age, sex	3343 cases (331, 9.9% CP, 2856, 85.4% CN, 156, 4.7% CC)	PCT (CN vs CP) CRP (CN vs CP) PCT (CC vs CP) CRP (CC vs CP)	0.76 0.64 0.86 0.65	79.8 - - -	56.2 - - -	
[10]	2013	prediction	Y	N	regression	1-3	suPAR, PCT, IL-6, CRP, WBC and combinations	132 patients (55, 41.7% CP)	suPAR (>7.9 ng/mL) PCT (>0.78 ng/mL) IL-6 (>493 pg/mL) CRP () WBC () suPAR PCT IL-6 PCT IL-6 SuPAR IL-6	0.73 0.74 0.74 0.60 0.57 0.80 0.79 0.78	- - - - - - - -	- - - - - - - -	
[11]	2014	prediction	Y	N	regression	7	vital signs (3), biomarkers (4)	2422 patients (14, 5.7% CP)	derivation internal validation external validation	0.78 0.75 0.79	- - -	- - -	
[12]	2015	prediction	Y	Y	regression tree-based SVM	52†	time-series data 24h before admission from which the median, mean and max are extracted. biomarkers (12x3=36) vital signs (4x3=16) combined (36+16=52)	2915 patients (control) 521 patients (target)	regression (biomarkers) tree-based (biomarkers) SVM (biomarkers) regression (vital signs) tree-based (vital signs) SVM (vital signs) regression (combined) tree-based (combined) SVM (combined)	0.83 0.86 0.84 0.76 0.76 0.84 0.88 0.88 0.87	0.55 0.93 0.61 0.44 0.44 0.59 0.60 0.61 0.64	0.92 0.59 0.93 0.90 0.91 0.94 0.93 0.93 0.94	
[13]	2016	prediction	Y	Y	score	17†	vital signs (5), biomarkers (3) demographics (age)	1394 patients (159, 11.4% CP)	validation	0.92	0.90	0.81	It needs time series data from 5 hours before admission. From each of these time series the mean and the difference between the first (Tn ₅) and the last (Tn) observations are used to create two vectors with 8 features. The MIMIC-II database was used.
[14]	2016	Prediction	Y	Y	score	17†	vital signs (6), Glasgow Comma Score, demographics (age)	17274 patients (1840, 10.65% CP)	InSight (0h onset time) InSight (4h after onset time) MEWS SOFA SIRS SAPS II	0.88 0.74 0.80 0.73 0.61 0.70	0.80 0.80 0.70 0.80 0.72 0.75	0.80 0.54 0.77 0.48 0.44 0.52	It needs time series data from 2 hours before admission. From each of these time series the mean and the difference between the first (Tn ₅) and the last (Tn) observations are used to create two vectors with 8 features. The MIMIC-III database was used.
[15]	2017	prediction	Y	Y	regression Gaussian tree-based linear SVM	10-12	symptoms (2), vital signs (6), demographics (2), free text from chief complaint and nurse assessment	230936 episodes (32103, 13.9% CP)	Vitals LSVM Vitals LR Vitals GNB Vitals RFC	0.67 0.67 0.65 0.70	0.56 - - -	0.68 - - -	

keys: CPN=causal probabilistic networks; CRP=c-reactive protein; GNB=gaussian naïve bayes; IL6=interleukin 6; LBP=lipopolysaccharide binding protein; LR=linear regression; LSVM=linear support vector machine; MEWS=modified early warning score; PCT=procalcitonin; RFC=random forest classifier; SAPS II=simplified acute physiology score; SIRS=systemic inflammatory response syndrome; SOFA=sequential organ failure assessment; suPAR=soluble urokinase-type plasminogen activator receptor; SVM=support vector machine; WBC=white blood cell counts;

† indicates those studies in which the input features were time series.

Table B.5: Literature review: diagnosis of bloodstream infection (II).

Ref	Year	Purpose	Inclusion c.	Sepsis c.	Method	n	Variables	Data	Outcomes	ROC	SENS	SPEC	Comments
[16]	2018	prediction	N	N	tree-based	30†	vital signs (6)	UCSF hospital: 90353 patients sepsis: 1179 (1.3%) severe sepsis: 349 (0.39%) sepsis shock: 614 (0.68%) MIMIC III: 21604 patients sepsis: 413 (1.91%) severe sepsis: 609 (2.82%) sepsis shock: 942 (4.36%)	InSight (sepsis)	0.92	0.80	0.95	It needs time series data from 7 hours before admission. The measurements are collected hourly. To construct the final feature vector, measurements for the first 3 hours (3) and corresponding differences of consecutive measurements (2) were used to create a final vector of 30 features (5 values x 6 variables).
									MEWS (sepsis)	0.76	--	--	
									SOFA (sepsis)	0.63	--	--	
									SIR (sepsis)	0.75	--	--	
									Insight (severe)	0.87	0.80	0.84	
	2018	prediction	N	N	GNB DTC RFC SVM ANN	6	biomarkers (6)	1236398 profiles (3636; 0.29%)	MEWS (severe)	0.77	--	--	The data from University College of San Francisco (UCSF) hospital and the MIMIC-III database were used.
									SOFA (severe)	0.65	--	--	
									SIR (severe)	0.72	--	--	
									Insight (shock)	0.99	0.80	0.99	
									MEWS (shock)	0.94	--	--	
	2018	prediction	N	N	GNB DTC RFC SVM ANN	6	biomarkers (6)	1236398 profiles (3636; 0.29%)	SOFA (shock)	0.86	--	--	The outliers were discarded using the interquartile (IQRx1.5) rule. For the best performing algorithms (SVM and ANN) results without discarding outliers were quite similar.
									SIR (shock)	0.82	--	--	
									GNB	0.95	0.83	0.91	
									DTC	0.94	0.82	0.95	
									RFC	0.95	0.82	0.97	
									SVM	0.96	0.86	0.93	
									ANN	0.96	0.87	0.97	

keys: ANN=artificial neural network; DTC=decision tree classifier; GNB=gaussian naïve bayes; MEWS=modified early warning score; RFC=random forest classifier; SIRS=systemic inflammatory response syndrome; SOFA=sequential organ failure assessment; SVM=support vector machine;
† indicates those studies in which the input features were time series.

Table B.6: Literature review: diagnosis of urinary tract infection.

Ref	Year	Purpose	Inclusion c.	UTI c.	Method	n	Variables	Data	Outcomes	ROC	SENS	SPEC	Comments
[17]	2018	prediction	Y	Y	regression	1	procalcitonin	293 patients (148, 50.5% CP, 155, 49.5% CN)	PCT (0.25ng/ml)	0.72	0.67	0.63	
[18]	2007	Identification of UTI predictors	Y	Y	ANN	30	symptoms (9), demographics (age) biomarkers (11), patient history (5), bacteria, others (3)	212 women patients Group 1: 56, 26.4% CP Group 2: 117, 55.19% CP	Group 1 Group 2	0.85 0.79	- -	- -	The study suggests that the presence of bacteria is a good predictor for urinary tract infections.
[19]	2014	APN prediction	Y	Y	regression	1	PCT, CRP, WCC within 25h of admission	46 patients (16 patients APN, 3 patients lower UTI)	APN PCT (1ng/ml) APN CRP (20mg/l) APN WBC (15000/mm3)	0.96 0.86 0.59	0.91 0.86 0.57	0.88 0.48 0.44	
[20]	2018	prediction	Y	Y	XGBoost RFC AdaBoost SVM elastic Net ANN LR	Full: 211 Reduced: 10	Reduced Symtoms (1) demographics (2), biomarkers (4), history of UTI, blood culture, bacteria	80387 visits & 55365 patients (18284, 28% CP)	full XGBoost	0.90	0.62	0.95	The superiority of XGBoost for UTI prediction was statistically significant in comparison to all other algorithms but RFC.
									full RFC	0.90	0.57	0.96	
									full AdaBoost	0.88	0.62	0.92	
									full SVM	0.88	0.50	0.97	
									full elastic Net	0.89	0.57	0.95	
									full LR	0.89	0.58	0.95	
									full ANN	0.89	0.55	0.95	
									red XGBoost	0.88	0.55	0.95	
									red RFC	0.86	0.55	0.94	
									red AdaBoost	0.83	0.62	0.89	
									red SVM	0.82	0.50	0.96	
									red elastic Net	0.87	0.52	0.95	
									red ANN	0.87	0.53	0.95	
									red LR	0.87	0.54	0.95	
	2018	prediction	N	positive urine culture	GNB DTC RFC SVM ANN	6	biomarkers (6)	1244784 profiles (12023; 0.97%)	GNB	0.82	0.59	0.87	The outliers were discarded using the interquartile (IQRx1.5) rule. For the best performing algorithms (SVM and ANN) results without discarding outliers were quite similar.
									DTC	0.76	0.47	0.95	
									RFC	0.78	0.45	0.98	
									SVM	0.83	0.64	0.86	
									ANN	0.82	0.66	0.83	

keys: ANN=artificial neural network; APN=acute pyelonephritis; CRP=c-reactive protein; DTC=decision tree classifier; GNB=gaussian naïve bayes; PCT=procalcitonin; RFC=random forest classifier; SVM=support vector machine; UTI=urinary tract infection; WBC=white blood cell counts; XGBoost=extreme gradient boost;

Table B.7: Literature review: diagnosis of respiratory tract infection.

Ref	Year	Purpose	Inclusion c.	RTI c.	Method	n	Variables	Data	Outcomes	ROC	SENS	SPEC	Comments
[21]	2017	predict nosocomial LRTI	Y	N	logistic regression (LR) fisher discriminant (FDA) risk index (RI)	11	symptoms (6), demographics (2), tracheotomy, prophylaxis antibiotic, disease	49328 cases (893, 1.8% CP)	internal validation (LR) internal validation (FDA) internal validation (RI) 10-fold CV (LR) 10-fold CV (FDA) 10-fold CV (RI)	0.91 - - 0.89 - 0.91	0.86 - - 0.87 - 0.87	0.79 - - 0.74 - 0.79	
[22]	2016	predict pathogen	N	N	tree-based	20	Model 1 (3): age, gender, season Model 2: mining from text Model 3: combination M1&M2	1685 patients	M1 Adenovirus M1 Influenza M1 Metapneumovirus M1 Parainfluenza M1 Rhinovirus M1 RSV M2 Adenovirus M2 Influenza M2 Metapneumovirus M2 Parainfluenza M2 Rhinovirus M2 RSV M3 Adenovirus M3 Influenza M3 Metapneumovirus M3 Parainfluenza M3 Rhinovirus M3 RSV	0.57 0.74 0.69 0.66 0.71 0.55 0.48 0.45 0.51 0.66 0.47 0.47 0.53 0.72 0.69 0.72 0.68 0.57	- - - - - - - - - - - - 0.97 1.00 0.97 0.95 0.97 0.94	- - - - - - - - - - - - 0.04 0.46 0.20 0.20 0.21 0.08	
[23]	2007	predict LRTI	N	N	logistic regression	1	vital signs (4) biomarkers (2)	364 patients (pulmonary malignancy 5, LRTI pneumonia 48, LRTI non-pneumonia 316)	hr>100bpm rr>22bpm o2sat <95% bt>38°C crp > 20mg/lx leukocyte count>=10million/ml	- - - - - -	0.25 0.50 0.52 0.30 0.73 0.80	0.90 0.79 0.80 0.89 0.65 0.46	Keys: o2sat=oxygen saturation, crp=c-reactive protein, hr=heart rate, rr=respiratory rate, bt=body temperature
[24]	2017	predict ARI	N	N	SVM tree-based (RFC) regression	Not specified	demographics (sex, age), symptoms (3), laboratory data, microbiology	Year 2013-2014: 414 signups Year 2014-2015: 623 signups	SVM RFC LR	- - -	- - 0.59	- 0.62 -	
[25]	2017	prediction of pneumonia in LRTI	Y	N	Regression	Not specified	age, gender, social deprivation, vital signs and medical history, and symptoms if present	28883 patients	bt bt + crackles bt + crackles + o2sat bt + crackles + o2sat + pulse	0.59 0.65 0.67 0.68	0.25 - - -	0.92 - - -	Keys: bt=body temperature, o2sat=oxygen saturation
[26]	2018	chest disease diagnosis	N	N	SVM ASVM	38	symptoms (7), vital signs (1), demographics (2), biomarkers (>20), microbiology (1), smoking addictions, others	357 samples TB: 50 COPD: 71 pneumonia: 60 asthma: 44 lung Cancer: 32 normal: 100	Accuracy	-	-	-	The study reports accuracy instead of area under the ROC, sensitivity and specificity. The accuracy results (%) for SVM ASVM were: TB: 92 94 COPD: 91 93 Pneumonia 90 92 Asthma: 97 99 Lung cancer: 97 90 Normal: 92 98
[27]	2018	chest disease diagnosis	N	N	deep learning	420	X-ray image	ChestX-ray14 dataset: training (28744 patients, 98637 images), validation (1672 patients, 6351 images) test (389 patients, 420 images).	Atelectasis Cardiomegaly Effusion Infiltration Mass Nodule Pneumonia	0.81 0.92 0.86 0.78 0.87 0.78 0.77	- - - - - - -	- - - - - - -	
	2018	prediction	N	positive blood culture	GNB DTC RFC SVM ANN	6	biomarkers (6)	1236732 profiles (3971; 0.32%)	GNB DTC RFC SVM ANN	0.90 0.87 0.90 0.92 0.93	0.73 0.65 0.65 0.81 0.83	0.90 0.94 0.96 0.92 0.87	The study discarded outliers applying the inter-quartile (IQRx1.5) rule. For the best performing algorithms (which were SVM and ANN) results without discarding outliers were quite similar.

keys: ANN=artificial neural network; ARI=acute respiratory tract infection; ASVM=A support vector machine; COPD=artificial neural network;DTC=decision tree classifier; FDA=fisher discriminant analysis;GNB=gaussian naïve bayes; LR=linear regression; LRTI=lower respiratory tract infection;RFC=random forest classifier; RI=risk index; RSV=respiratory syncytial virus; RTI=respiratory tract infection; SVM=support vector machine; TB=tuberculosis;

Table B.8: Literature review: diagnosis of surgical site infection.

Ref	Year	Purpose	Inclusion c.	SSI c.	Method	n	Variables	Data	Outcomes	ROC	SENS	SPEC	Comments
[28]	2016	Prediction	Y	N	m-regression	17	demographics (3), patient history (2), microbiology, other	2364 patients (131 overall SSI, 33 incisional SSI, 98 organ/space SSI)	Logistic R Scoring (BBOT)	0.74 0.74			
[29]	2018	Prediction	N	N	logistic regression multivariate logistic regression ANN	LR pre-operative: 14 LR post-operative: 20 ANN pre-operative: 29 ANN post-operative: 50	patient epidemiologic data (3), disease characteristic (5), operative data (3), demographics (2), biomarkers (9), others	1854 reconstructions (438 presented SSI)	LR pre-operative LR post-operative ANN pre-operative	0.69 0.72 0.81	0.14 0.22 0.61	0.95 0.93 0.89	
[30]	2015	Prediction	N	NSQI gold standard	multivariate logistic regression	Over 150	preoperative characteristics, intraoperative factors, postoperative morbidity occurrences	year 1 – 6258 procedures (504, 8.05% CP) year 2 – 3996 procedures (278, 6.95% CP) year 3 – 2262 procedures (127, 5.6% CP)	Superficial Deep Organ/space total	0.82 0.90 0.89 0.90	0.57 0.21 0.36 0.78	0.90 0.99 0.97 0.87	The study did not present the sensitivity. The values were calculated afterwards based on the description of the data.
[31]	2012	Prediction	N	N	Tree based	9	intraoperative factors, and postoperative characteristics	Training 35551 (144, 0.32% CP) DH 1429 (77, 5.4% CP) IH 11586 (713, 6.15% CP) VA SLCHCS 568 (35, 6.2% CP) VVMC 852 (17, 2% CP)	DH IH VA SLCHCS VVMC		0.46 0.36 0.50 0.00	0.95 0.94 0.94 0.98	
	2018	prediction	N	positive wound culture	GNB DTC RFC SVM ANN	6	biomarkers (6)	1237630 profiles (4869; 0.39%)	GNB DTC RFC SVM ANN	0.90 0.85 0.88 0.92 0.92	0.73 0.62 0.62 0.78 0.79	0.92 0.98 0.98 0.91 0.80	The outliers were discarded using the interquartile (IQRx1.5) rule. For the best performing algorithms (SVM and ANN) results without discarding outliers were quite similar.

keys: SVM=support vector machine; SSI=surgical site infection; NSQI=lower respiratory tract infection; RFC=random forest classifier; DTC=decision tree classifier; ANN=artificial neural network; LR=linear regression; DH=denver health; IH=intermountain health; VA SLCHCS=veterans affairs salt lake city health care system; VVMC=vail valley medical centre;

Table B.9: Literature review: diagnosis of culture positive

Ref	Year	Purpose	Inclusion c.	Sepsis c.	Method	n	Variables	Data	Outcomes	ROC	SENS	SPEC	Comments
[32]	2017	prediction	Y	N	Youden's index	1	procalcitonin	519 patients Pulmonary: 223 (43%) Bacteremia: 21 (4%) Urinary: 67 (13%) Wound: 26 (5%) <i>Clostridium. Difficile</i> : 10 (2%) gastrointestinal (not <i>C.diff</i>): 67 (13%) other: 104 (20%)	All cultures Pulmonary Urinary Bacteremia	0.62 0.49 0.43 0.78			It evaluates whether initial procalcitonin values can determine the positive of any culture. The inclusion criteria is: age>18, infection as admitted diagnosis, bacterial culture and PCT value within 24h. Patients were excluded if they were receiving immunosuppressive therapy (see antibiotics below) and/or had HIV or cancer. Antibiotics: belowmercaptopurine, prednisone, antithymocyte, globulin, azathioprine, basiliximan, cyclophosphamide, cyclosporine, calcineurin inhibitors, everolimus, methotrexate, mycophenolate, rituximab and sirolium
	2018	prediction	N	N	GNB DTC RFC SVM ANN	6	biomarkers (6)	1268301 profiles (41524; 3.27% CP)	GNB DTC RFC SVM ANN	0.89 0.88 0.89 0.89 0.91	0.72 0.65 0.65 0.77 0.78	0.89 0.95 0.96 0.87 0.88	The outliers were discarded using the interquartile (IQRx1.5) rule. For the best performing algorithms (SVM and ANN) results without discarding outliers were quite similar.

keys: ANN=artificial neural network; APN=acute pyelonephritis; CRP=c-reactive protein; DTC=decision tree classifier; GNB=gaussian naïve bayes; PCT=procalcitonin; RFC=random forest classifier; SVM=support vector machine; UTI=urinary tract infection; WBC=white blood cell counts; XGBoost=extreme gradient boost;

BIBLIOGRAPHY

- [1] HJ Gómez-Vallejo, B Uriel-Latorre, M Sande-Meijide, B Villamarín-Bello, Reyes Pavón, F Fdez-Riverola, et al. "A case-based reasoning system for aiding detection and classification of nosocomial infections". *Decision Support Systems* 84 (2016), pp. 104–116.
- [2] Mariana Maceiras Cabrera and Ernesto Ocampo Edye. "Integration of rule based expert systems and case based reasoning in an acute bacterial meningitis clinical decision support system". *arXiv preprint arXiv:1003.1493* (2010).
- [3] Mehdi Neshat, Mehdi Sargolzaei, Adel Nadjaran Toosi, and Azra Masoumi. "Hepatitis disease diagnosis using hybrid case based reasoning and particle swarm optimization". *ISRN Artificial Intelligence* 2012 (2012).
- [4] Nassim Douali, Huszka Csaba, Jos De Roo, Elpiniki I Papageorgiou, and Marie-Christine Jaulent. "Diagnosis support system based on clinical guidelines: comparison between case-based fuzzy cognitive maps and Bayesian networks". *Computer methods and programs in biomedicine* 113.1 (2014), pp. 133–143.
- [5] SL Ting, Siu Keung Kwok, Albert HC Tsang, and WB Lee. "CASESIAN: A knowledge based system using statistical and experiential perspectives for improving the knowledge sharing in the medical prescription process". *Expert Systems with Applications* 37.7 (2010), pp. 5336–5346.
- [6] Mical Paul, Steen Andreassen, Anders D Nielsen, Evelina Tacconelli, Nadja Al-manasreh, Abigail Fraser, et al. "Prediction of bacteremia using TREAT, a computerized decision-support system". *Clinical Infectious Diseases* 42.9 (2006), pp. 1274–1282.
- [7] Nathan I Shapiro, Richard E Wolfe, Sharon B Wright, Richard Moore, and David W Bates. "Who needs a blood culture? A prospectively derived and validated prediction rule". *The Journal of emergency medicine* 35.3 (2008), pp. 255–264.
- [8] M Tromp, Benno Lansdorp, CP Bleeker-Rovers, JM Klein Gunnewiek, BJ Kullberg, and P Pickkers. "Serial and panel analyses of biomarkers do not improve the prediction of bacteremia compared to one procalcitonin measurement". *Journal of infection* 65.4 (2012), pp. 292–301.
- [9] Seri Jeong, Yongjung Park, Yonggeun Cho, and Hyon-Suk Kim. "Diagnostic utilities of procalcitonin and C-reactive protein for the prediction of bacteremia determined by blood culture". *Clinica Chimica Acta* 413.21-22 (2012), pp. 1731–1736.

- [10] Martin Hoenigl, Reinhard B Raggam, Jasmin Wagner, Thomas Valentin, Eva Leitner, Katharina Seeber, et al. "Diagnostic accuracy of soluble urokinase plasminogen activator receptor (suPAR) for prediction of bacteremia in patients with systemic inflammatory response syndrome". *Clinical biochemistry* 46.3 (2013), pp. 225–229.
- [11] Jungyoup Lee, Seung Sik Hwang, Kyuseok Kim, You Hwan Jo, Jae Hyuk Lee, Joonghee Kim, et al. "Bacteremia prediction model using a common clinical test in patients with community-acquired pneumonia". *The American journal of emergency medicine* 32.7 (2014), pp. 700–704.
- [12] Joseph Guillén, Jiankun Liu, Margaret Furr, Tianyao Wang, Stephen Strong, Christopher C Moore, et al. "Predictive models for severe sepsis in adult ICU patients". *Systems and Information Engineering Design Symposium (SIEDS)*, 2015. IEEE. 2015, pp. 182–187.
- [13] Jacob S Calvert, Daniel A Price, Uli K Chettipally, Christopher W Barton, Mitchell D Feldman, Jana L Hoffman, et al. "A computational approach to early sepsis detection". *Computers in biology and medicine* 74 (2016), pp. 69–73.
- [14] Thomas Desautels, Jacob Calvert, Jana Hoffman, Melissa Jay, Yaniv Kerem, Lisa Shieh, et al. "Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach". *JMIR medical informatics* 4.3 (2016).
- [15] Steven Horng, David A Sontag, Yoni Halpern, Yacine Jernite, Nathan I Shapiro, and Larry A Nathanson. "Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning". *PloS one* 12.4 (2017), e0174708.
- [16] Qingqing Mao, Melissa Jay, Jana L Hoffman, Jacob Calvert, Christopher Barton, David Shimabukuro, et al. "Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU". *BMJ open* 8.1 (2018), e017833.
- [17] Alexander R Levine, Midori Tran, Jonathan Shepherd, and Edgar Naut. "Utility of initial procalcitonin values to predict urinary tract infection". *The American journal of emergency medicine* (2018).
- [18] Paul S Heckerling, Gay J Canaris, Stephen D Flach, Thomas G Tape, Robert S Wigton, and Ben S Gerber. "Predictors of urinary tract infection based on artificial neural networks and genetic algorithms". *International Journal of Medical Informatics* 76.4 (2007), pp. 289–296.
- [19] Rui-Ying Xu, Hua-Wei Liu, Ji-Ling Liu, and Jun-Hua Dong. "Procalcitonin and C-reactive protein in urinary tract infection diagnosis". *BMC urology* 14.1 (2014), p. 45.
- [20] R Andrew Taylor, Christopher L Moore, Kei-Hoi Cheung, and Cynthia Brandt. "Predicting urinary tract infections in the emergency department with machine learning". *PloS one* 13.3 (2018), e0194085.
- [21] Yong Chen, Xue Shan, Jingya Zhao, Xuelin Han, Shuguang Tian, Fangyan Chen, et al. "Predicting nosocomial lower respiratory tract infections by a risk index based system". *Scientific reports* 7.1 (2017), p. 15933.

- [22] Mark V Mai and Michael Krauthammer. "Controlling testing volume for respiratory viruses using machine learning and text mining". *AMIA Annual Symposium Proceedings*. Vol. 2016. American Medical Informatics Association. 2016, p. 1910.
- [23] Anette Holm, Joergen Nexoe, Lene A Bistrup, Svend S Pedersen, Niels Obel, Lars P Nielsen, et al. "Aetiology and prediction of pneumonia in lower respiratory tract infection in primary care". *Br J Gen Pract* 57.540 (2007), pp. 547–554.
- [24] Bisakha Ray and Rumi Chunara. "Predicting Acute Respiratory Infections from Participatory Data". *Online journal of public health informatics* 9.1 (2017).
- [25] Michael Moore, Beth Stuart, Paul Little, Sue Smith, Matthew J Thompson, Kyle Knox, et al. "Predictors of pneumonia in lower respiratory tract infections: 3C prospective cough complication cohort study". *European Respiratory Journal* 50.5 (2017), p. 1700434.
- [26] Amani Yahyaoui and Nejat Yumusak. "Decision support system based on the support vector machines and the adaptive support vector machines algorithm for solving chest disease diagnosis problems". *Biomedical Research* 29.7 (Apr. 2018), pp. 1474–1480.
- [27] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, et al. "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning". *arXiv preprint arXiv:1711.05225* (2017).
- [28] Ru-Hong Tu, Chang-Ming Huang, Jian-Xian Lin, Qi-Yue Chen, Chao-Hui Zheng, Ping Li, et al. "A scoring system to predict the risk of organ/space surgical site infections after laparoscopic gastrectomy for gastric cancer based on a large-scale retrospective study". *Surgical endoscopy* 30.7 (2016), pp. 3026–3034.
- [29] Pao-Jen Kuo, Shao-Chun Wu, Peng-Chen Chien, Shu-Shya Chang, Cheng-Shyuan Rau, Hsueh-Ling Tai, et al. "Artificial neural network approach to predict surgical site infection after free-flap reconstruction in patients receiving surgery for head and neck cancer". *Oncotarget* 9.17 (2018), p. 13768.
- [30] Zhen Hu, Gyorgy J Simon, Elliot G Arsoniadis, Yan Wang, Mary R Kwaan, and Genevieve B Melton. "Automated detection of postoperative surgical site infections using supervised methods with electronic health record data". *Studies in health technology and informatics* 216 (2015), p. 706.
- [31] Agency for Healthcare Research. *Improving the Measurement of Surgical Site Infection Risk Stratification/Outcome Detection*. Tech. rep.
- [32] Erica M Caffarini, Joshua DeMott, Gourang Patel, and Ishaq Lat. "Determining the clinical utility of an absolute procalcitonin value for predicting a positive culture result". *Antimicrobial agents and chemotherapy* (2017), AAC–02007.



DOCUMENTATION

The server side of EPiC IMPOC processes queries, interacts with the permanent storage and provides decision support. It has been implemented in Java and uses an object relational mapping java library (Hibernate ORM) to map an object-oriented domain model to a traditional relational database (SQL). The implementation follows the Representational State Transfer (REST) architectural design. The functionality of the system has been divided into the following sections: CBR case (see Table C.1), patient (see Table C.3), doctor (see Table C.2), authentication (see Table C.4), inference (see Table C.5) and others (see Table C.6). These are briefly explained below.

Table C.1: EPiC IMPOC interface: cases.

CBR CASE	
/case/add	add a new case to the database
/case/copy	clone an existing case from the database
/case/update	update the parameters of an existing case within the database
/case/delete	delete a case from the database
/case/list	list all the cases in the database
/case/retrieve	get a case from the database
/case/report	create a report of the case in a pdf document
/case/pathology	get the pathology laboratory markers associated to a case (time series)
/case/pathology/last	get the last pathology laboratory markers associated to a case
/case/pathology/dates	get the pathology markers between the specified dates
/case/sensitivity	get the microbiology results associated to a case
/case/sensitivity/dates	get the microbiology results between the specified dates

Table C.2: EPiC IMPOC interface: doctors.

DOCTOR	
/doctor/add	add a new doctor to the database
/doctor/list	list all the doctors in the database

Table C.3: EPiC IMPOC interface: patients.

PATIENT	
/patient/add	add a new patient to the database.
/patient/delete	delete a patient from the database
/patient/list	list all the patients in the database
/patient/admitted	lists patients that have an open case in the database
/patient/admitted/ward	list patients that have an open case from an specific ward
/patient/retrieve	get a patient from the database
/patient/update	update the information of an existing patient within the database

Table C.4: EPiC IMPOC interface: authentication.

AUTHENTICATION	
/auth/login	authenticate the user
/auth/logout	logout the user

Table C.5: EPiC IMPOC interface: inference.

INFERENCE	
/inference/infection/features	get features used for inference
/inference/infection/case	perform inference of positive culture for a case
/inference/culture/case	perform inference of positive cultures on specific sites for a case

Table C.6: EPiC IMPOC interface: others.

OTHERS	
/bugs/list	list all the pathogens in the database
/drug/add	add an antimicrobial to the database
/drug/update	update an antimicrobial to the database
/drug/list	list the antimicrobials in the database

D

LIST OF ACRONYMS: ORGANISMS AND ANTIMICROBIALS

This section presents the list of acronyms for the organisms (see Table D.1) and antimicrobials (see Table D.2). The first table, which corresponds to organisms, presents the code (acronym), the name (including genus and species), the genus and the Gram status. The second table, which corresponds to antimicrobials, presents the code (acronym), the name and the category. Note that those codes starting with “A_” were generated automatically since no code was provided within the dataset provided by Imperial College London NHS Trust.

Table D.1: Antibiotics: notation and description.

Organism Code	Organism Name	Genus	Gram Status
A_ABOTROP	Abiotrophia species	abiotrophia	
A_ACTINOMY	Actinomyces sp	actinomyces	p
A_AFLAVUS	Aspergillus flavus	aspergillus	
A_AFUMIGATU	Aspergillus fumigatus	aspergillus	
A_AGNB	Anaerobic Gram Negative Bacillus	bacillus	p
A_AGPB	Anaerobic Gram Positive Bacillus	bacillus	p
A_AHAEMOLYT	Acinetobacter haemolyticus	acinetobacter	
A_ACHAEMOLYT	Arcanobacterium haemolyticum	arcanobacterium	
A_AISRAELII	Actinomyces israelii	actinomyces	p
A_ALCALIGE	Alcaligenes sp	alcaligenes	
A_ALTERNAR	Alternaria sp	alternaria	
A_AMEYERI	Actinomyces meyeri	actinomyces	p
A_ANAEROBE	Anaerobes		
A_ANAESLUND	Actinomyces naeslundii	actinomyces	p
A_ANIGER	Aspergillus niger	aspergillus	
A_AODONTOLY	Actinomyces odontolyticus	actinomyces	p
A_ARADIORES	Acinetobacter radioresistens	acinetobacter	
A_ASCHAALII	Actinobaculum schaalii	actinobaculum	
A_ASTREPTOC	Anaerobic Streptococcus	streptococcus	p
A_AUSTUS	Aspergillus ustus	aspergillus	
A_BCIRCULAN	Bacillus circulans	bacillus	p
A_BIFIDOBA	Bifidobacterium sp.	bifidobacterium	
A_BMELITENS	Brucella melitensis	brucella	
A_BORDETEL	Bordetella sp.	bordetella	
A_BREVUNDI	Brevundimonas sp.	brevundimonas	
A_BSGD	Beta-haemolytic strep. Group D	streptococcus	p
A_BSGF	Beta-haemolytic strep. Group F	streptococcus	p
A_BSNAG	Beta-haemolytic strep. NOT A-G	streptococcus	p
A_BSUBTILIS	Bacillus subtilis	bacillus	p
A_BURKHOLD	Burkholderia sp.	burkholderia	n
A_BVESICULA	Brevundimonas vesicularis	brevundimonas	
A_CALBIDUS	Cryptococcus albidus	cryptococcus	
A_CANDIDA	Candida sp.	candida	
A_CBOVIS	Corynebacterium bovis	corynebacterium	p
A_CDUBLINIE	Candida dubliniensis	candida	
A_CFAMATA	Candida famata	candida	
A_CGUILLIER	Candida guilliermondii	candida	
A_CHOMINIS	Cardiobacterium hominis	cardiobacterium	
A_CHRYSEOB	Chryseobacterium sp.	chryseobacterium	
A_CHRYSEOM	Chryseomonas sp.	chryseomonas	
A_CLADOSPO	Cladosporium sp.	cladosporium	
A_CLAURENTI	Cryptococcus laurentii	cryptococcus	
A_CLIPOLYTI	Candida lipolytica	candida	
A_CLUSITANI	Candida lusitaniae	candida	
A_CMENINGOS	Chryseobacterium meningosepticum	chryseobacterium	
A_CMINUTISS	Corynebacterium minutissimum	corynebacterium	p
A_CNEOFORMA	Cryptococcus neoformans	cryptococcus	

Table D.1: Antibiotics: notation and description.

Organism Code	Organism Name	Genus	Gram Status
A_CPAPARSIL	Candida parapsilosis	candida	
A_CPELLICUL	Candida pelliculosa	candida	
A_CPSEUDODI	Corynebacterium pseudodiphtheriticum	corynebacterium	p
A_CRYPTOCO	Cryptococcus sp.	cryptococcus	
A_CSEPTICUM	Clostridium septicum	clostridium	p
A_CSTREPTOC	Carboxyphilic Streptococcus	streptococcus	p
A_CTESTOSTE	Comamonas testosteroni	comamonas	
A_CTUBERCUL	Corynebacterium tuberculostearicum	corynebacterium	p
A_CUNIGUTTU	Cryptococcus uniguttulatus	cryptococcus	
A_CUTILIS	Candida utilis	candida	
A_EAMNIGENU	Enterobacter amnigenus	enterobacter	n
A_EAVIUM	Enterococcus avium	enterococcus	p
A_ECOLI	escherichia coli	escherichia	n
A_EDURANS	Enterococcus durans	enterococcus	p
A_EGERGOVIA	Enterobacter gergoviae	enterobacter	n
A_EHERMANII	Escherichia hermanii	escherichia	n
A_EIKENELL	Eikenella sp.	eikenella	
A_ELUDWIGII	Enterobacter ludwigii	enterobacter	n
A_ENTEROBA	enterobacter sp.	enterobacter	n
A_ESAKAZAKI	Enterobacter sakazakii	enterobacter	n
A_ESCHERIC	Escherichia sp.	escherichia	n
A_FFUNGI	Filamentous fungi		
A_FLAVOBAC	Flavobacterium sp.	flavobacterium	
A_FNI	Fungi NOT isolated		
A_FNUCLEATU	Fusobacterium nucleatum	fusobacterium	n
A_FOXYSPORU	Fusarium oxysporum	fusarium	
A_FUSARIUM	Fusarium sp.	fusarium	
A_FUSOBACT	Fusobacterium sp.	fusobacterium	n
A_GCANDIDUM	Geotrichum candidum	geotrichum	
A_GCAPITATU	Geotrichum capitatum	geotrichum	
A_GEOTRICH	Geotrichum species	geotrichum	
A_GMORBILLO	Gemella morbillorum	gemella	
A_GVAGINALI	Gardnerella vaginalis		
A_HAFNIA S	Hafnia sp.	hafnia	
A_HAPHROPHI	Haemophilus aphrophilus	haemophilus	
A_KKINGAE	Kingella kingae		
A_KLOECKER	Kloeckera species	kloeckera	
A_KPLANTICO	Klebsiella planticola	klebsiella	n
A_LACTOBAC	LactoBacillus sp.	lactobacillus	p
A_LEUCONOS	Leuconostoc sp.	leuconostoc	
A_LFERMENTU	Lactobacillus fermentum	lactobacillus	p
A_LISTERIA	Listeria sp.	listeria	p
A_LPARACASE	Lactobacillus paracasei	lactobacillus	p
A_LRHAMNOSU	Lactobacillus rhamnosus	lactobacillus	p
A_LWADEI	Leptotrichia wadei		
A_MABSCUSSU	Mycobacterium abscessus	mycobacterium	p
A_MAVIUM	Mycobacterium avium	mycobacterium	p
A_MBOVIS	Mycobacterium bovis	mycobacterium	p
A_MCHELONAE	Mycobacterium chelonae	mycobacterium	p
A_MFORTUITU	Mycobacterium fortuitum	mycobacterium	p
A_MICROBAC	Microbacterium sp.	microbacterium	
A_MINTRACEL	Mycobacterium intracellulare	mycobacterium	p
A_MKANSASII	Mycobacterium kansasii	mycobacterium	p
A_MMALMOENS	Mycobacterium malmoense	mycobacterium	p
A_MMARINUM	Mycobacterium marinum	mycobacterium	p
A_MMUCOGENI	Mycobacterium mucogenicum	mycobacterium	p
A_MPEREGRIN	Mycobacterium peregrinum	mycobacterium	p
A_MSCROFULA	Mycobacterium scrofulaceum	mycobacterium	p
A_MSIMAE	Mycobacterium simae	mycobacterium	p
A_MSTREPTOC	Microaerophilic streptococcus	streptococcus	p
A_MTC	Mycobacterium tuberculosis complex	mycobacterium	p
A_MTUBERCUL	Mycobacterium tuberculosis	mycobacterium	p
A_MXENOPI	Mycobacterium xenopi	mycobacterium	p
A_MYCOBACT	Mycobacterium sp.	mycobacterium	p
A_MYCOPLAS	Mycoplasma sp.	mycoplasma	p
A_NCINEREA	Neisseria cinerea	neisseria	n
A_NMUCOSA	Neisseria mucosa	neisseria	n
A_OANTHROPI	Ochrobactrum anthropi	ochrobactrum	
A_OCHROBAC	Ochrobactrum sp.	ochrobactrum	
A_PALCALIGE	Pseudomonas alcaligenes	pseudomonas	n
A_PASTEURE	Pasteurella sp.	pasteurella	
A_PENICILL	Penicillium sp.	penicillium	
A_PPTOSTR	Peptostreptococcus sp.	streptococcus	p
A_PFLUORESC	Pseudomonas fluorescens	pseudomonas	n
A_PLILACINU	Paecilomyces lilacinus		

Table D.1: Antibiotics: notation and description.

Organism Code	Organism Name	Genus	Gram Status
A_PMEANINO	Prevotella melaninogenica	prevotella	
A_PPROPIONI	Propionibacterium propionicus	propionibacterium	p
A_PREVOTEL	Prevotella sp.	prevotella	
A_PROVIDEN	Providencia sp.	providencia	
A_PVARIOTII	Paecilomyces variotii		
A_RAOULTEL	Raoultella sp.	raoultella	
A_RARRHIZUS	Rhizopus arrhizus	rhizopus	
A_RHIZOPUS	Rhizopus sp.	rhizopus	
A_RHODOCOC	Rhodococcus sp.	rhodococcus	
A_RHODOTOR	Rhodotorula sp.	rhodotorula	
A_RMINUTA	Rhodotorula minuta	rhodotorula	
A_RMUCILAGI	Rhodotorula mucilaginosa	rhodotorula	
A_RPICKETTI	Ralstonia pickettii		
A_RTERRIGEN	Raoultella terrigena	raoultella	
A_SAPIOSPER	Scedosporium apiospermum	scedosporium	
A_SBOVIS	Streptococcus bovis	streptococcus	p
A_SBOYDII	Shigella boydii	shigella	n
A_SCERVISIA	Saccharomyces cervisiae		
A_SDYSENTER	Shigella dysenteriae	shigella	n
A_SMUTANS	Streptococcus mutans	streptococcus	p
A_SNEWPORT	Salmonella newport	salmonella	n
A_SODORIFER	Serratia odorifera	serratia	n
A_SPA	Salmonella paratyphi A	salmonella	n
A_SPAUCIMOB	Sphingomonas paucimobilis	sphingomonas	
A_SPHINGOM	Sphingomonas sp.	sphingomonas	
A_SPORCINUS	Streptococcus porcinus	streptococcus	p
A_SPOROTRI	Sporotrichum sp.	sporotrichum	
A_STYPHI	Salmonella typhi	salmonella	n
A_STYPHIMUR	Salmonella typhimurium	salmonella	n
A_SVESTIBUL	Streptococcus vestibularis	streptococcus	p
A_SVIRIDANS	Streptococcus viridans	streptococcus	p
A_TASAHII	Trichosporon asahii	trichosporon	
A_TINKIN	Trichosporon inkin	trichosporon	
A_TMUCOIDES	Trichosporon mucoides	trichosporon	
A_VALGINOLY	Vibrio alginolyticus	vibrio	n
A_VIBRIO S	Vibrio sp.	vibrio	n
A_VPARAHAEM	Vibrio parahaemolyticus	vibrio	n
A_YERSINIA	Yersinia sp.	yersinia	n
A_YPSEUDOTU	Yersinia pseudotuberculosis	yersinia	n
ABAU	Acinetobacter baumannii	acinetobacter	
ACAV	Aeromonas caviae	aeromonas	
ACHRO	Achromobacter sp.	achromobacter	
ACINE	Acinetobacter sp.	acinetobacter	
AEROC	Aerococcus sp.	aerococcus	
AEROM	Aeromonas sp.	aeromonas	
AERU	Aerococcus urinae	aerococcus	
AFAE	Alcaligenes faecalis	alcaligenes	
AHS	Alpha haemolytic Streptococcus	streptococcus	p
AHYD	Aeromonas hydrophila	aeromonas	
ALWO	Acinetobacter lwoffii	acinetobacter	
AVIR	Aerococcus viridans	aerococcus	
AXYL	Achromobacter xylosoxidans	achromobacter	
BACIL	Bacillus sp.	bacillus	p
BACTE	Bacteroides sp.	bacteroides	
BCEP	Burkholderia cepacia	burkholderia	n
BCER	Bacillus cereus	bacillus	p
BDIM	Brevundimonas diminuta	brevundimonas	
BFRA	Bacteroides fragilis	bacteroides	
BHS	Beta-haemolytic streptococcus	streptococcus	p
BHSA	Beta-haemolytic strep. Group A	streptococcus	p
BHSB	Beta-haemolytic strep. Group B	streptococcus	p
BHSB	Beta-haemolytic strep. Group B	streptococcus	p
BHSC	Beta-haemolytic strep. Group C	streptococcus	p
BHSCG	Beta-haemolytic Strep. Group C/G	streptococcus	p
BHSG	Beta-haemolytic strep. Group G	streptococcus	p
BOVA	Bacteroides ovatus	bacteroides	
BPSE	Burkholderia pseudomallei	burkholderia	n
BREVI	Brevibacterium sp.	brevibacterium	
BVUL	Bacteroides vulgatus	bacteroides	
CALB	Candida albicans	candida	
CAMA	Citrobacter amalonaticus	citrobacter	
CAMPY	Campylobacter sp.	campylobacter	n
CBRA	Citrobacter braakii	citrobacter	
CCOL	Campylobacter coli	campylobacter	n
CFET	Campylobacter fetus	campylobacter	n

Table D.1: Antibiotics: notation and description.

Organism Code	Organism Name	Genus	Gram Status
CFRE	Citrobacter freundii	citrobacter	
CGLA	Candida (Torulopsis) glabrata	candida	
CHAE	Corynebacterium haemolyticum	corynebacterium	p
CIND	Chryseobacterium indologenes	chryseobacterium	
CITRO	Citrobacter sp.	citrobacter	
CJEI	Corynebacterium jeikeium	corynebacterium	p
CJEJ	Campylobacter jejuni	campylobacter	n
CKEF	Candida kefyr	candida	
CKOS	Citrobacter koseri	citrobacter	
CKRU	Candida krusei	candida	
CLOST	Clostridium sp.	clostridium	p
CNS	Coagulase negative staphylococcus	staphylococcus	p
COLIF	Coliform sp.	coliform	
CORYN	Corynebacterium sp	corynebacterium	p
CPER	Clostridium perfringens	clostridium	p
CSTR	Corynebacterium striatum	corynebacterium	p
CTRO	Candida tropicalis	candida	
DACI	Delftia acidovorans	delftia	
DIPHT	Diphtheroids		
EASB	Enterobacter asburiae	enterobacter	n
ECASS	Enterococcus casseliflavus	enterococcus	p
ECLO	Enterobacter cloacae	enterobacter	n
ECOL	Escherichia coli	escherichia	n
ECOR	Eikenella corrodens	eikenella	
EFAM	Enterococcus faecium	enterococcus	p
EFAS	Enterococcus faecalis	enterococcus	p
EGAL	Enterococcus gallinarum	enterococcus	p
EMEN	Elizabethkingia meningoseptica	elizabethkingia	
ENTAE	Enterobacter aerogenes	enterobacter	n
ENTB	Enterobacter sp.	enterobacter	n
ENTC	Enterococcus sp.	enterococcus	p
FNEC	Fusobacterium necrophorum	fusobacterium	n
GEM	Gemella sp.	gemella	
GHAEM	Gemella haemolysans	gemella	
HAEMO	Haemophilus sp.	haemophilus	
HALV	Hafnia alvei	hafnia	
HHA	Haemophilus haemolyticus	haemophilus	
HINF	Haemophilus influenzae	haemophilus	
HPAP	Haemophilus paraphrophilus	haemophilus	
HPHA	Haemophilus parahaemolyticus	haemophilus	
HPIN	Haemophilus parainfluenzae	haemophilus	
HPYL	Helicobacter pylori		
KLEBS	Klebsiella sp.	klebsiella	n
KLUYV	Kluyvera sp.	kluyvera	
KOXY	Klebsiella oxytoca	klebsiella	n
KPNE	Klebsiella pneumoniae	klebsiella	n
LFC	Lactose fermenting coliform	coliform	
LMON	Listeria monocytogenes	listeria	p
MCAT	Moraxella catarrhalis	moraxella	
MICROC	Micrococcus sp.	micrococcus	
MMOR	Morganella morganii	morganella	
MODO	Myroides odoratimimus		
MORAX	Moraxella sp.	moraxella	
NEISS	Neisseria sp.	neisseria	n
NFLA	Neisseria flavescens	neisseria	n
NGON	Neisseria gonorrhoeae	neisseria	n
NHS	Non-haemolytic streptococcus	streptococcus	p
NLF	Non-lactose Fermenting Coliform	coliform	
NMEN	Neisseria meningitidis	neisseria	n
NOCAR	Nocardia sp.	nocardia	
PACN	Propionibacterium acnes	propionibacterium	p
PAENI	Paenibacillus sp.	paenibacillus	
PAER	Pseudomonas aeruginosa	pseudomonas	n
PANA	Peptostreptococcus anaerobius	streptococcus	p
PANTO	Pantoea sp.	pantoea	
PBIV	Prevotella bivia	prevotella	
PLUT	Pseudomonas luteola	pseudomonas	n
PMIR	Proteus mirabilis	proteus	n
PMUL	Pasteurella multocida	pasteurella	
PORI	Pseudomonas oryzae	pseudomonas	n
PPUT	Pseudomonas putida	pseudomonas	n
PRET	Providencia rettgeri	providencia	
PROPI	Propionibacterium sp.	propionibacterium	p
PROTE	Proteus sp.	proteus	n
PSEUD	Pseudomonas sp.	pseudomonas	n

Table D.1: Antibiotics: notation and description.

Organism Code	Organism Name	Genus	Gram Status
PSHI	Plesiomonas shigelloides		
PSTU	Pseudomonas stutzeri	pseudomonas	n
PSTUA	Providencia stuartii	providencia	
PVUL	Proteus vulgaris	proteus	n
RMUC	Rothia mucilaginosa	rothia	
RORN	Raoultella ornithinolytica	raoultella	
ROTH	Rothia sp.	rothia	
RRAD	Rhizobium radiobacter		
SAGA	Streptococcus agalactiae	streptococcus	p
SALMO	Salmonella sp.	salmonella	n
SANG	Streptococcus anginosus	streptococcus	p
SAUR	Staphylococcus aureus	staphylococcus	p
SCON	Streptococcus constellatus	streptococcus	p
SDYSEQ	Streptococcus dysgalactiae ssp equisimilis	streptococcus	p
SEPI	Staphylococcus epidermidis	staphylococcus	p
SEQU	Streptococcus equi	streptococcus	p
SERRA	Serratia sp.	serratia	n
SFLE	Shigella flexneri	shigella	n
SGOR	Streptococcus gordonii	streptococcus	p
SHAEM	Staphylococcus haemolyticus	staphylococcus	p
SHIGE	Shigella sp.	shigella	n
SINT	Streptococcus intermedius	streptococcus	p
SLIQ	Serratia liquifaciens	serratia	n
SLUG	Staphylococcus lugdunensis	staphylococcus	p
SMAL	Stenotrophomonas maltophilia	strenotrophomonas	n
SMAR	Serratia marcescens	serratia	n
SMIL	Streptococcus milleri group	streptococcus	p
SMIT	Streptococcus mitis	streptococcus	p
SORA	Streptococcus oralis	streptococcus	p
SPARAS	Streptococcus parasanguis	streptococcus	p
SPNE	Streptococcus pneumoniae	streptococcus	p
SPYO	Streptococcus pyogenes	streptococcus	p
SSAL	Streptococcus salivarius	streptococcus	p
SSAN	Streptococcus sanguis	streptococcus	p
SSAP	Staphylococcus saprophyticus	staphylococcus	p
SSON	Shigella sonnei	shigella	n
STAPH	Staphylococcus sp.	staphylococcus	p
STREP	Streptococcus sp.	streptococcus	p
VEILL	Veillonella sp.	veillonella	
VIRST	Viridans streptococcus	streptococcus	p
YEAST	Yeasts		
YENT	Yersinia enterocolitica	yersinia	n

Table D.2: Antimicrobials: notation and description.

Antibiotic Code	Antibiotic Name	Category
AAMPC	amp c markers	
AAMI	amikacin	aminoglycosides
AAMO	amoxycillin	aminopenicillins
AAMPH	amphotericin	
AAND	anidulafungin	
AAUG	augmentin	
AAZI	azithromycin	macrolides
AAZT	aztreonam	monobactams
ABAC	bacitracin	polypeptides
ACPO	cpo marker	
A_CAPREOMY	capreomycin	
ACAS	caspofungin	
A_CEFEPIME	cefepime	cephalosporins
ACIX	cefixime	cephalosporins
ACTX	cefotaxime	cephalosporins
ACXT	cefoxitin	cephalosporins
A_CEFPODOX	cefpodoxime	cephalosporins
ACAZ	ceftazidime	cephalosporins
ACONE	ceftriaxone	cephalosporins
ACXM	cefuroxime	cephalosporins
ACELX	cephalexin	cephalosporins
A_CEPHAZOL	cephazolin	cephalosporins
ACHL	chloramphenicol	
ACIP	ciprofloxacin	fluoroquinolones
ACLA	clarithromycin	macrolides
ACLI	clindamycin	macrolides
A_CLOTRIMO	clotrimazole	
ACOL	colistin sulphate	polypeptides
ACOT	cotrimoxazole	sulfonamides
ADAP	daptomycin	
A_DOXYCYCL	doxycycline	tetracyclines
AERT	ertapenem	meropenems
AESBL	esbl markers	
AERY	erythromycin	macrolides
A_ETHAMBUT	ethambutol	
AMET	flucloxacillin	penicillins
AFLUZ	fluconazole	
AFLUC	flucytosine	
AFOS	fosfomycin	
AFUS	fusidic acid	
AGEN	gentamicin	aminoglycosides
AGEN	gentamicin 200	aminoglycosides
AIMP	imipenem	meropenems
A_ISONIAZI	isoniazid	
AITR	itraconazole	
A_KANAMYCI	kanamycin	aminoglycosides
A_KETOCONA	ketoconazole	
ALEV	levofloxacin	fluoroquinolones
ALIN	linezolid	oxazolidinones
AMLS	mls markers	
AMEC	mecillinam	penicillins
AMER	meropenem	meropenems
AMTZ	metronidazole	nitroimidazoles
AMF	micafungin	
A_MICONAZO	miconazole	
A_MINOCYCL	minocycline	tetracyclines
AMOX	moxifloxacin	fluoroquinolones
AMUP	mupirocin	
ANAL	naladixic acid	fluoroquinolones
ANEO	neomycin	aminoglycosides
ANIT	nitrofurantoin	
ANOV	novobiocin	aminocoumarin
AOFL	ofloxacin	fluoroquinolones
AOPT	optochin	
AOXA	oxacillin	penicillins
APEF	pefloxacin	
APEN	penicillin	penicillins
APCZ	posaconazole	
A_PROTHION	prothionamide	
A_PYRAZINA	pyrazinamide	
A_RIFABUTI	rifabutin	
ARIF	rifampicin	
ASEP	septrin	
ASPE	spectinomycin	aminoglycosides
A_STREPTOM	streptomycin	aminoglycosides

Table D.2: Antimicrobials: notation and description.

Antibiotic Code	Antibiotic Name	Category
A_SULBACTA	sulbactam	sulfonamides
A_SULPHAME	sulphamethoxazole	
ASYN	synercid	
ATAZ	tazocin	glycopeptide
ATEI	teicoplanin	
ATEM	temocillin	
A_TERBINAF	terbinafine	penicillins
ATET	tetracycline	
ATIG	tigecycline	tetracyclines
A_TIMENTIN	timentin	aminoglycosides
ATOB	tobramycin	
ATRI	trimethoprim	
AVAN	vancomycin	glycopeptide
AVOR	voriconazole	

DEFINITION OF THE ATTRIBUTES

This section describes the vital symptoms, demographics, laboratory tests and other information included in the CBR case for the antimicrobial therapy selection clinical trial performed in the ICU (see Table E.1). First, the attributes have been classified in four groups: binary (B), numeric (N), categorical (C) and set (S). For each of the attributes the range and description are included. Note that the range represents the possible values that can be assigned to the attribute. To conclude, the table also defines the distance/similarity metric of the attribute within the CBR.

Table E.1: Feature attributes in the CBR case description.

	Attribute	Similarity	Values	Description
B ₁	hiv	S _{equal}	True, False	The patient has HIV
B ₂	pregnant	S _{equal}	True, False	The patient is pregnant
B ₃	diabetes	S _{equal}	True, False	The patient has diabetes
B ₄	catheter	S _{equal}	True, False	The patient has a catheter
B ₅	central line	S _{equal}	True, False	The patient has a central line
N ₁	age	S _{interval}	[16, 105]	Age of the patient
N ₂	height	S _{interval}	[130, 230]	Height of the patient (cm).
N ₃	weight	S _{interval}	[35, 150]	Weight of the patient (kg)
N ₄	heart rate	S _{interval}	[0, 200]	Heart rate of the patient
N ₅	body temperature	S _{interval}	[35, 41]	The body temperature (celsius)
N ₆	respiratory rate	S _{interval}	[14, 40]	Respiratory rate
N ₇	blood pressure high	S _{interval}	[0, 200]	High blood pressure (mm of mercury)
N ₈	blood pressure low	S _{interval}	[0, 200]	Low blood pressure (mm of mercury)
N ₉	O ₂ saturation	S _{interval}	[0, 100]	Fraction of oxygen in the blood
N ₁₀	ionotropes	S _{interval}	-	Agent that alters muscular contractions
N ₁₁	lactate	S _{interval}	[0, 10]	Lactate levels (mg/dL)
C ₁	gender	S _{equal}	-	None, Male or Female
C ₂	renal support	S _{equal}	-	None, HDX, CVVH
C ₃	chest examination	S _{table}	-	None, Crackles, Dull, Clear
C ₄	chest radiography	S _{table}	-	None, Consolidation, Effusion, Clear, AUD
C ₅	abd. examination	S _{table}	-	None, Peritonitis, Tender, SNT
S ₁	allergies	None	-	The allergies of the patient
S ₂	organs infected	S _{set}	-	The organs under suspicion of infection
S ₃	pathogens	S _{cluster}	-	The pathogens reported from microbiology

Keys: B_i=boolean attribute; N_i=numeric attribute; C_i=categorical attribute; S_i=set attribute; O₂=oxygen; HIV=human immunodeficiency virus; cm=centimeters; kg=kilograms; mg/dL=milligram per deciliter; HDX=expanded hemodialysis; CVVH=continuous veno-venous hemofiltration; AUD=air under diaphragm; HIV=soft not tender;