Understanding *Clostridium difficile* Infection Outcomes, through Host Clinical Variables, and Bacterial Whole Genome and Phenotypic Analysis

A thesis submitted by Emma Butt to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Biological Sciences
In February 2013

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University

Abstract

Clostridium difficile is a clinically problematic pathogen and continues to persist within the healthcare system. Presentation of disease symptoms ranges from mild to severe diarrhoea, through to fulminant pseudomembranous colitis. Approximately 20% of patients will suffer from recurrent episodes and of all patients who die from *C. difficile* related causes, approximately 41% of death certificates mention *C. difficile* as an underlying cause of death, and this poses a significant burden on healthcare facilities.

Three methods of investigation were employed to develop a more comprehensive understanding of both the host and isolate association with the outcomes of *C. difficile* infection; mortality and recurrence. These methods were; analysing patient clinical data to try and identify host markers of infection outcomes, evaluating *C. difficile* type association with infection outcomes, and genetically and phenotypically characterising the clinically relevant *C. difficile* isolates associated to these outcomes.

During this study statistical analysis of clinical data revealed that there were four variables; white cell count, serum albumin, C-reactive protein and respiratory rate, which were prognostic of mortality in patients with *C. difficile* infection. Threshold levels of these variables were used to create a clinical prediction rule to classify patients with *C. difficile* infection who were more 'at risk' from mortality, with statistical significance in both a derivation and validation cohort. However, analysis was unable to determine variables prognostic of recurrent infection.

Due to small sample sizes of some groups of isolates, no groups of *C. difficile* isolates were significantly associated with increased recurrent infection or mortality during this study. Some groups of isolates were associated with primary only infection and/or low mortality. There was a non-significant trend in particular *C. difficile* isolate groups being associated with infection outcomes; a panel of representative isolates was therefore chosen to be characterised in more detail.

Phenotypic and genetic analysis of a panel of sixteen *C. difficile* isolates revealed isolate specific differences in toxin production, conservation of transposable elements and SNP abundances, which may have played a role in infection outcomes. Isolate motility and antibiotic resistance profiles were not statistically significantly different between isolates within a particular group of *C. difficile* types.

One hypothesis from the collective results obtained during this study suggests that the phenotypic and genotypic changes in isolates may have facilitated differences in their interaction with the host. In turn, the host specific inflammatory response to the infecting *C. difficile* isolate may have played a role in host outcomes.

Research conducted during this study has begun to assess which host specific responses may be important in determining the outcome of *C. difficile* infection, and which *C. difficile* isolate characteristics may in part also contribute. However, the assessment of both host and isolate association to infection outcomes would benefit from further investigation in a larger cohort, in order to prove or refute conclusively any hypothesises generated in this study.

Acknowledgements

I would first and foremost like to dedicate this PhD thesis to my granddad, Alec (John) Ireland; he never got to see me finish the research on which this PhD thesis is based. As a great chemist, his enthusiasm for science was bestowed upon me, and without which, I am sure I would not have pursued a career in science with such passion as I now do.

You are always in my thoughts granddad.....

To all those who made this PhD possible I wish to sincerely thank you. I am sincerely grateful to my principle supervisor Dr Steve Michell, your continued support has been invaluable to me. It has allowed me to develop as scientist and as a result of this, have found a research area which I now wish to pursue to a higher level.

Thanks to my second supervisor Dr Ed Keedwell, who has provided much needed statistical support and whose patience was greatly appreciated.

To everyone on the 4th floor of Geoffrey Pope, I am extremely grateful for your help and guidance in the lab and for making a great working environment at the University of Exeter. I have thoroughly enjoyed working with you all.

Finally, I would like to thank my parents.... You let me move back home to do this PhD and for that I am eternally grateful. I have loved every minute of living back home with you! My fiancée Jamie..... You are the light when the moments are dark, you have been patient, kind and have accepted all that has come with this PhD. I love you so much and am very lucky to have you in my life.

Table of Contents

Content **Page Number**

Page Number

Page Number

Page Number

Page Number

Page Number

Page Number

Page Number

List of Figures and Tables **Page Number**

Figures

Chapter One

Chapter Two

Chapter Three

Chapter Four

Page Number

Chapter Five

Chapter Six

Page Number

Chapter Seven

Chapter Eight

Page Number

Page Number

Page Number

Tables **Page Number**

Chapter One

Chapter Two

Chapter Three

Chapter Four

Page Number

Page Number

Chapter Nine

Concluding Discussion

Chapter Ten **Page Number**

Appendices

List of Manuscripts Submitted/Prepared as a Result of Research Conducted during this PhD

Submitted Manuscripts

Derivation and Validation of a Simple, Accurate and Robust Prediction Rule for Risk of Mortality in Patients with *Clostridium difficile* Infection. Emma Butt[1], Jane AH Foster[2], Edward Keedwell[3], Julia EA Bell[4], Richard W Titball[1], Aneel Bhangu[5], Stephen L Michell[1] and Ray Sheridan[6]. *BMC Infectious diseases. Under review*

***Clostridium difficile*: Evaluation of disease severity and the potential role of Faecal Calprotectin as a biomarker. Have the UK guidelines got it right?** Dr Jane AH Foster[1], Emma Butt[2], Dr Edward Keedwell[3], Dr Stephen L Michell[2], Dr Tariq Ahmed[1], Dr Ray Sheridan[1]. *In preparation.*

Characterisation of Clinically Relevant *Clostridium difficile* Isolates from different Infection Outcomes. [1]Emma Butt, [2]Ed Keedwell, and [1]Stephen .L. Michell. *Journal of Medical Microbiology. In preparation.*

Abbreviations

ADP	Adenine tri phosphate
H_1	Alternative hypothesis
AUC	Area under the curve
bp	Base pairs
bpm	Beats per minute
BHI	Brain heart infusion
mcL	Cells/micro litre
χ^2	Chi-Square value
CDI	*Clostridium difficile* infection
CDS	Coding Sequences
Cfu/s	Colony forming units
df	Degrees of freedom
DNA	Deoxyribonucleic acid
ELISA	Enzyme-linked immunosorbent assay
g	Grams
x g	Gravitational force
Galß1-4GlcNAc	N-Acetyllactosamine
GDH	Glutamate Dehydrogenase
GDP	Guanosine 5'-diphosphate
GTP	Guanosine 5'-Triphosphate
HPA	Health Protection Agency
HCAI	Healthcare acquired infection
hr/hrs	hour/hours
IL	Interleukin
Kbp	Kilo base pairs
kg	Kilo gram
L	Litres
MAP	Mitogen Activated Protein
µl	Micro litres
µM	Micro molar
mg	Milli grams
ml	Milli litres
mm	Milli metre
mM	Milli molar
MIC	Minimum inhibitory concentration
Mins	Minutes
MLST	Multi locus sequence typing
mAb/s	Monoclonal Antibody/ies
Nf-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
ng	Nano grams
nm	Nano metre

NGS	Next generation sequencing
NS	Non-synonymous
H_0	Null hypothesis
PaLoc	Pathogenicity locus
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
ROC	Receiver operating curves
REA	Restriction enzyme analysis
Rpm	Revolutions per minute
SEM	Standard error of the mean
S	Synonymous
Tcdx	*Clostridium difficile* Toxin (x)= A, B, C, R, or E
Tn	Transposon
TE	Tris- EDTA
TAE	Tris-acetate EDTA
TBS	Tris-buffered saline
TBST	Tris-buffered saline
TY	Tryptone yeast
TNF	Tumour necrosis factor
Yrs	Years