Developing individual and population-level models of hospital-acquired Clostridium difficile infection (CDI) to enable the timely identification of high-risk patients, to facilitate inter-institution comparisons of incidence, and to promote quality improvement

By

Kevin Antoine Brown

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy in Epidemiology

Dalla Lana School of Public Health, in the University of Toronto

© Copyright by Kevin Antoine Brown 2014
Abstract

*Clostridium difficile* infection (CDI) is a health-care associated disease that causes diarrhea and, in more severe cases, colitis and death. It is the most common of hospital-acquired infections in North America. In this thesis, I present 3 studies elucidating individual and population-level risk factors for CDI.

In a meta-analysis of antibiotic effects on community-associated CDI risk, I found that Clindamycin, fluoroquinolones, and cephalosporins, monobactams and carbapenems (CMCs) had the largest effects on CDI risk, while macrolides, sulfonamides and trimethoprim and penicillins had lesser associations with CDI risk. I noted no effect of tetracyclines on CDI risk.

A retrospective case-cohort analysis of 2,067 adults hospitalized at a tertiary hospital in Ontario, Canada, between June 2010 and May 2012 was conducted to assess the magnitude and duration of antibiotic effects on CDI risk. I found that risks due to antibiotic exposure begin two days after the initiation of antibiotics and are most elevated for a period of 5 days after antibiotic cessation. In this cohort study, exposure to symptomatic CDI cases was not associated with increased risk of infection.

I carried out a time-series analysis on 16 years of monthly CDI incidence rate data in the United States to describe the relationship between CDI incidence and hospital pneumonia and influenza (P&I) prevalence at a hospital network level. Peak P&I
prevalence preceded peak CDI incidence by 9 weeks and I found that surges in hospital P&I prevalence were associated with disproportionately large impacts on CDI incidence that lasted for 13 months.

These findings suggest that control of CDI must involve a strengthening of antimicrobial stewardship initiatives in order to diminish inappropriate antibiotic exposures. Initiatives that would lessen the impacts of seasonal respiratory infections on hospital admissions, including influenza vaccination and programs to reduce pneumonia-associated hospitalizations, could also help stem the increasing incidence of CDI in the North American health care system.
Acknowledgements

As anyone who’s completed a doctorate knows, it takes a village to bring a thesis into the world. I would never have been able to finish my thesis without the steadfast support of my advisors, collaborators, friends, and family.

I would like to begin by thanking my thesis supervisor, Dr. David Fisman, for being an incredible source of encouragement and for mobilizing my ideas from day one. In particular, I would like to thank him for teaching me how to foster great collaborative relationships through purposefulness, encouragement, and enthusiasm. I am also grateful to the other members of my advisory committee who helped propel my projects forward with their insights, availability, and can-do attitude. I am indebted to Dr. Nick Daneman for being the paragon of humility and reliability, for providing me with amazing and rapid feedback, and for always helping me understand the clinical implications of my research. I wish to give a special thanks to Dr. Allison McGeer, for her amazing consistency and humility, and the breadth of her expertise that she brought to each of my committee meetings. And lastly, I would to thank Dr. Rahim Moineddin, for lively discussions about multilevel and time-series statistical analyses at his University avenue office.

I wish to thank all the professors who have inspired and helped me along the way. In particular, Drs. Amy Greer, Sue Bondy, Laura Rosella, Nancy Krieger, Dionne Gesink, John McLaughlin and Michael Escobar.

I also wish to thank my collaborators over the last several years. Especially, I would like to thank Paul Arora, for his wise mentorship and assistance throughout the entire PhD process. I am grateful to Kim Valenta, for stirring discussions about gut bacteria and ecological statistics, and her practical help revising sections of this thesis. I am especially grateful to Dr. Laurie Lambert for her inspiring dedication to the goal of improving cardiology services in Quebec and for her patience with me while I worked on this thesis.

My thesis would not have been possible without the assistance of several amazing collaborators and friends in France. I will always be grateful to Drs. Nicolas Voirin and Philippe Vanhems for hosting me at the Hôpital Édouard Hérriot in Lyon. I also wish to express my thanks to Nagham Khanafer, for being an amazing friend and collaborator and for propelling me into several of the most productive months of my research career thus far. Also, I would like to thank Vasily Aushev and Julie Camonis, for opening their homes to me while I was visiting in Lyon.

Many thanks are extended to my classmates at the University of Toronto for continual support and great discussions both inside and outside the classroom. In particular, I would like to thank Kindra St. Jacques, for bringing a critical mass to our graduate student space and Jessica Dennis for warming up the 7th floor of the Public Health School building. Also, I wish to thank Wallis Rudnick for being a great co-organizer and positive force in our program. I would like to say thank you to Emma Richardson for our chance encounters that made Toronto seem that much homier.

Thank you to my friends in Toronto for accompanying me through the highs and lows of the last several years. To Stephen Ross and Jane Hudson, for punting and bringing corn bread into my life. To Angel Chen, for being a steadfast dinner party partner and for commiserating with me. To Maya Goldenberg, for helping me engage in the most essential distractions that are eating, dancing and picnicking, for finding me two great homes in Toronto, and for connecting me with countless friends. To Matt Cassils, for sauerkraut, skiing and suds. To Chloé Camonis, for eating cheese on mountaintops, for a most amazing
companionship, and for a great tie. To my climbing friends Moti Yona, Vinh Truong, Alex Tevlin, Levi Goertz, Sophie Barbier, Hedi Buchner and Sebastian Hegelmier for putting up with the climbing bully that I am. To my fermentation friends John O, Amy Jiang, Steve Kwok and Alexander Gopen for enduring my hand waving and its sometimes disastrous results. I would also like to thank all of my Hot Yam and Harvest Noon friends for providing an instant community of like-minded foodies.

Finally, I would like to thank my family for making it all possible. I am grateful to my brother, Erik, for making the trip to Toronto multiple times in the 4 years I was here. And of course my dearest thanks go to my mother and father for their love and support on this doctoral journey.
Table of Contents

Abstract ................................................................................................................................. ii
Acknowledgements .............................................................................................................. iv
Table of Contents .................................................................................................................. vi
List of Tables ......................................................................................................................... ix
List of Figures ....................................................................................................................... x
Abbreviations ....................................................................................................................... xii

Chapter 1. Introduction .......................................................................................................... 1

1.1 Background .................................................................................................................. 1
  1.1.1 Microbiology of Clostridium difficile ................................................................. 2
  1.1.2 Population-Level Trends of CDI Incidence ....................................................... 3
  1.1.3 Community-associated CDI .............................................................................. 4
  1.1.4 Transmission of CDI ......................................................................................... 5
  1.1.5 Quantifying Risk of Transmission: Disease Pressure ....................................... 7
  1.1.6 Antibiotic-Associated Risk and the Role of Gut Flora .................................... 8
  1.1.7 Clostridium difficile Seasonality ..................................................................... 10
  1.1.8 Other CDI Risk Factors ................................................................................... 11
  1.1.9 Statistical Analysis of Hospital-Acquired Infections ....................................... 12

1.2 Note on the Format of the Thesis .............................................................................. 14

1.3 Thesis Objectives and their Scientific Contribution ............................................... 14
  1.3.1 Objective 1: Systematic Review on Antibiotics and Community-Associated CDI 15
  1.3.2 Objective 2: Duration of CDI Risk Associated with Antibiotic Exposure ....... 16
  1.3.3 Objective 3: Population-Level Seasonal CDI Variation .................................. 17

1.4 References .................................................................................................................... 17

Chapter 2. Antibiotics and the risk of community-associated Clostridium difficile infection (CDI): a meta-analysis ............................................................... 24

2.1 Reader’s Note .............................................................................................................. 24

2.2 Abstract ...................................................................................................................... 24

2.3 Introduction ............................................................................................................... 25

2.4 Methods ..................................................................................................................... 26
  2.4.1 Search Criteria ................................................................................................. 26
  2.4.2 Screening and Data Abstraction .................................................................... 27
  2.4.3 Quality Assessment ......................................................................................... 28
  2.4.4 Variables .......................................................................................................... 29
  2.4.5 Statistical Analysis .......................................................................................... 29

2.5 Results ......................................................................................................................... 30
  2.5.1 Pooled Effects .................................................................................................. 34
  2.5.2 Antibiotic Types ............................................................................................... 36
  2.5.3 Meta-Regression ............................................................................................... 36
  2.5.3 Publication Bias ............................................................................................... 37
  2.5.4 Antibiotic-Associated CDI Risk (AACR) Index ............................................. 37
  2.5.5 Secondary Analysis .......................................................................................... 38

2.6 Discussion ................................................................................................................... 38

2.7 Acknowledgements ..................................................................................................... 43
4.9 Appendix ................................................................................................................................. 100
  4.9.1 Bivariate Analyses ............................................................................................................. 100
  4.9.2 Design Effect ..................................................................................................................... 100
  4.9.3 Additional Figures ............................................................................................................. 101

Chapter 5. Discussion .................................................................................................................. 102
  5.1 Introduction ............................................................................................................................ 102
  5.2 Antibiotic Exposure ............................................................................................................... 102
  5.3 Age ........................................................................................................................................ 105
  5.4 Infection Pressure .................................................................................................................. 106
  5.5 Pneumonia and Influenza Effects and Seasonality .............................................................. 108
  5.6 Innovative Uses of Statistical Methods .............................................................................. 111
      5.6.1 Meta-Regression of CDI Risk Factors ............................................................................. 111
      5.6.2 Cohort Studies of Hospital Acquired Infections ............................................................. 112
      5.6.3 Co-Seasonality of CDI and P & I .................................................................................. 113
  5.7 Limitations ............................................................................................................................. 114
  5.8 Implications ............................................................................................................................ 116
  5.9 Potential Areas for Further Research and Methods Development .................................... 117
      5.9.1 Clostridium difficile Transmission Sources .................................................................. 117
      5.9.2 Impacts of CDI Testing Rates on CDI Incidence Measurement ................................. 118
      5.9.3 Prediction of CDI Spread in Health Care Networks ..................................................... 121
      5.9.4 Statistical Modeling Strategies and the Independent Censoring Assumption 124
  5.10 Conclusion ............................................................................................................................ 124
  5.11 Citations ................................................................................................................................ 125
List of Tables

Table 1.1: Thesis objectives and research questions ................................................................. 15
Table 2.1. Study quality assessment questions ........................................................................ 28
Table 2.2: Selected characteristics of eligible articles ............................................................. 32
Table 2.3: Quality characteristics of eligible studies ............................................................... 34
Table 3.1. Selected Characteristics of Case and Control Patients, Sunnybrook Hospital, Toronto, Canada, June 2010 to May 2012 ............................................................................. 57
Table 3.2. Timing and Magnitude of CDI Risk Associated with 4 Different Antibiotic Exposure Measures .................................................................................................................... 58
Table 3.3. CDI Risk Associated with Antimicrobial Exposures During and Within 5 days of the End of Antimicrobial Therapy, for Antibiotic Risk Indexes and Specific Antibiotic Exposures .......................................................................................................................... 59
Table 3.4. CDI Risk Associated with Exposures Retained in the Adjusted Analyses* ............ 69
Table 4.1: Incidence of CDI (per 1,000 discharges) in Hospitalized Inpatients, United States, 1993 to 2008* ................................................................................................................................. 88
Table 4.2: Harmonic Regression Model Estimates for Monthly P&I Prevalence and CDI Incidence, United States, 1993 to 2008 ........................................................................................................... 90
List of Figures

Figure 1.1: Proposed transmission routes relating shedding to disease acquisition for nosocomial pathogens (32) ................................................................. 7
Figure 1.2: Clostridium difficile infection seasonality in Quebec (21) .................. 10
Figure 1.3: Best fit distributed lag model for the effect of capital appropriations on expenditures in the manufacturing industry (55) ....................................... 14
Figure 1.4: DLNM estimates of the lagged effects of temperature on mortality (56) .... 14
Figure 2.1: Flow chart of studies screened and included ........................................ 31
Figure 2.2: Pooled and study-specific risk estimates of community-associated CDI risk by antibiotic class .................................................................................. 35
Figure 2.3: The linear association between a 4 point antibiotic risk index and community-associated CDI risks .................................................................................. 38
Figure 3.1. Patients and Patient-Time Included in the Final Cohort, Sunnybrook Hospital, Toronto, Canada, June 2010 to May 2012 ..................................................... 55
Figure 3.2. Symptom Onset Time of Incident Clostridium difficile Infection (N=127), Sunnybrook Hospital, Toronto, Canada, June 2010 to May 2012 .................. 56
Figure 3.3. The Magnitude and Duration of Clostridium difficile Infection Risk After Antibiotic Therapy, Sunnybrook Hospital, Toronto, Canada, June 2010 to May 2012. Among inpatients, the incidence of Clostridium difficile infection was highest in the period 3 to 14 days after the start of antibiotic therapy, during antibiotic therapy, and within 5 days of the end of antibiotic therapy. * Includes patients without any identified antibiotic use ...................................................................................... 60
Figure 4.1: Measurement of a difference in peak timing for two annual harmonic time series. For each time series, the timing of peak incidence or phase shift (w1, w2) is calculated using the coefficients of the sine and cosine terms. The phase shift difference (w2 - w1) can then be computed by subtraction. On the X axis, week number 0 and 52 correspond to the date of January 1st in consecutive years .................................................. 85
Figure 4.2: Composition of the study population, United States, 1993-2008. All n correspond to unweighted counts of sampled hospital discharges. NHDS, National Hospital Discharge Survey; LOS, length of stay; CDI, Clostridium difficile infection; P&I, pneumonia and influenza .................................................................................. 87
Figure 4.3: Annual timing of CDI peak relative to the pneumonia and influenza peak for 20 age-region strata, United States, 1993-2008. On the X axis, week 0 corresponds to the date of January 1st. CDI, Clostridium difficile infection .................................................. 91
Figure 4.4: Distributed lag model estimated effect of hospital P&I prevalence on current and future Clostridium difficile incidence, United States, 1993-2008. The unadjusted model effects (squares) represent the lag effects for the single, overall time series, while the adjusted effects (+) are the mean lag effects across all 20 age-region strata. Shaded regions correspond to the 95% confidence bands for the lag effects .................................................................................. 92
Figure 4.5: Observed CDI incidence and fitted values of distributed lag prediction model, United States, 1993-2008. A 3 month central moving average was applied to the observed incidence for this visualization. CDI, Clostridium difficile infection ........................................ 93
Figure 4.6: Observed CDI incidence and fitted values of harmonic prediction model, United States, 1993-2008. A 3 month central moving average was applied to the observed incidence for this visualization. CDI, Clostridium difficile infection. Dotted line: observed incidence. Solid line: predicted incidence ........................................................................... 101
Figure 5.1: Monthly case counts at a hospital in Oxfordshire, England (23) ................................108
Figure 5.2: Suggested pathways for the association between seasonal respiratory infections and CDI risk. .............................................................................................................................................110
Figure 5.3: Panel A. The association between *Clostridium difficile* testing rates and, in Panel A, reported CDI rates, in Panel B, the per-test positivity of *Clostridium difficile* test results. Data are from a secondary analysis of a study of 97 hospitals across 34 European countries and both regression lines are based on simple linear regression of the 28 countries providing testing, patient follow-up time, and incidence data (48). .............................................................................................................................................120
Figure 5.4: The effect of outbreaks on network and the effect of outbreaks on hospitals, according to hospital size (53) .............................................................................................................................................122
Abbreviations

CDI: *Clostridium difficile* infection

CI: confidence interval

ICD-9-CM: *International Classification of Diseases, Ninth Revision, Clinical Modification*

IR: incidence rate

IRR: incidence rate ratio

LOS: length of stay

MIC: minimum inhibitory concentration

NHDS: National Hospital Discharge Survey

OR: odds ratio

P&I: pneumonia and influenza.

RR: relative risk
Chapter 1. Introduction

1.1 Background

Clostridium difficile is a toxin-producing bacterium that can proliferate in the gut and cause diarrhea, colitis, and death. Clostridium difficile primarily infects elderly populations that are hospitalized or have had recent contact with the health care system (1).

Clostridium difficile infection (CDI) pathogenesis initially involves spore ingestion (2). Since spores, and to a lesser extent vegetative bacteria, are highly resistant they can survive high stomach acidity levels as they travel towards the gut (3). Germination to the vegetative state, wherein bacteria can reproduce, excrete toxins and induce overt disease in hosts, has been shown to be induced by bile salts (4,5). Current or recent antibiotic use means that gut flora may be unbalanced and this compromised state of the gut ecosystem permits proliferation of Clostridium difficile bacteria (6). Cytotoxins and enterotoxins are released by the bacteria, and resulting symptoms include colitis and diarrhea, which in turn leads to environmental contamination and further spread of the disease. One treatment option is the prescription of targeted antibiotics, most commonly metronidazole or oral vancomycin. Recurrence with antibiotic treatment is frequent, on the order of 25%. Treatment via fecal transplant is highly effective among recurrent cases, though it is still uncommonly practiced in hospitals (6).

The global purpose of this thesis is to clarify the role of specific classes of antibiotics in the acquisition of CDI, to better understand the timing of CDI acquisition with respect to antibiotic exposure, and to improve our understanding of the determinants of population-level CDI incidence.
1.1.1 Microbiology of *Clostridium difficile*

*Clostridium difficile* was first isolated from the feces of newborns in the 1930s (7). It is a Gram positive bacterium, produces disease-causing cytotoxins, and can live in a highly durable sporulated state.

*Clostridium difficile* produces 2 main toxins: enterotoxin A and cytotoxin B, which are encoded by genes tcdA and tcdB located within a 19.6-kb region of the *Clostridium difficile* genome known as the pathogenicity locus (PaLoc) (8). The main clinical symptoms of CDI can be explained by the effects of the toxins tcdA and tcdB. These toxins cause the disruption of the actin cytoskeleton and tight junctions, and result in decreased transepithelial resistance, fluid accumulation and destruction of the intestinal epithelium. A study of 50 toxigenic and 39 nontoxigenic *Clostridium difficile* isolates verified that the PaLoc was present in all 50 toxigenic strains, and absent in the 39 non-toxigenic strains (9). Recent research involving the removal of tcdA and tcdB from the *Clostridium difficile* genome showed that genetically manipulated tcdA-/tcdB+ strains produce disease amongst hamsters whereas tcdA+/tcdB- strains do not produce disease (10). This suggests that toxin B is the key virulence mechanism.

*Clostridium difficile* is capable of forming an endospore, which is a tough outer coating made up of DNA and part of the cell cytoplasm. The process of endospore formation is unclear for anaerobes (11). The spore form of *Clostridium difficile* is not responsible for disease manifestations, but is critical to understanding disease transmission and relapse. Specifically, *Clostridium difficile* spores can survive for long periods on hospital surfaces. Furthermore they are resistant to prolonged exposure at high temperatures (60°C and up to 3 hours at 70°C) and 70% ethanol, but not resistant to
sporicidal agents, such as 10% bleach (2). *Clostridium difficile* spores are resistant to high stomach acidity levels, and therefore facilitate fecal-oral transmission. Finally, spores are an important factor for disease relapse, since they are impermeable to antibiotics and germinate after treatment.

Recent increases in the incidence and virulence of CDI across North America and Europe have been attributed to the emergence of a new strain of *Clostridium difficile*, known as strain BI/NAP1/027. The deletion of the tcdC toxin repressor gene in the BI/NAP1/027 causes 20-fold increase in production of toxins A and B relative to other strains (12). A recent study has shown that two genetically distinct BI/NAP1/027 lineages acquired an identical mutation conferring a high level of fluoroquinolone resistance (13) in north-eastern North America beginning in 2001. These two independent lineages subsequently spread globally to the UK, continental Europe, Asia, and Australia. Mouse intoxication assays have shown that the particular type of B toxin produced by strain BI/NAP1/027 is 4 times more lethal than other type B toxins (14). An additional recent study has shown that the BI/NAP1/027 strain may be more capable of proliferating in typical human gut environments as evidenced by its capacity to outcompete other *Clostridium difficile* strains in *in vitro* CDI models (15).

### 1.1.2 Population-Level Trends of CDI Incidence

*Clostridium difficile* bacteria became widely recognized as a cause of clindamycin-associated colitis in the 1970s. Rapidly thereafter, bacterial culture methods for identifying *Clostridium difficile* were introduced (16,17). Since this baseline period, the reported incidence of hospital-acquired CDI has increased steadily as have the diagnostic methods capable of detecting it. In the United States, the incidence of CDI (17) within ICUs
of large hospitals increased by over 50% in the 1987 to 2001 period, from less than 3 cases per 10,000 patient days to over five cases per 10,000 patient days (18). In the same period, incidence increased even more rapidly in older patients (19). Since the early 2000s, the incidence of CDI in the United States has increased even more rapidly, from 3.82 per 1,000 discharges in 2000 to 8.75 per 1,000 discharges in 2008 (20).

This recent surge is arguably associated with the emergence of the new, more virulent strains of *Clostridium difficile* (1). Many studies have considered the link between increasing CDI incidence and the spread of hypervirulent strains (21,22). These studies have demonstrated that inter- and intra-national increases in CDI incidence may be associated with increases in NAP1 strain prevalence. Beginning in October 2002 and peaking in March 2004, hospitals across the province of Quebec, Canada experienced a three-fold increase in CDI incidence (21). During this province-wide outbreak, over 80% of cases in one hospital were infected with the hypervirulent BI/NAP1/027 strain. In a study of hospitals from 34 European countries, a high fractional prevalence of NAP1 strain was associated with higher CDI incidence at the national-level (22).

### 1.1.3 Community-associated CDI

Cases of CDI having no apparent link to a health care facility make up almost half of CDI cases (23,24) but are far less frequently studied in the literature. Patients with community-associated disease are likely to have been exposed to health care settings as outpatients, to infants less than 1 year of age, or to household members with active CDI which suggests that health care settings drive community-associated disease (25). Other (non-environmental) risk factors for community-associated CDI are similar to those for
hospital-associated disease and include antibiotic exposure and proton-pump inhibitor exposure (24).

1.1.4 Transmission of CDI

_Clostridium difficile_ is transmitted by the fecal-oral route, wherein spores or vegetative bacteria from a _Clostridium difficile_ source are ingested by a susceptible patient, via direct person-to-person contact, via a vector, or via airborne dissemination. Suspected sources include both symptomatic and asymptomatically colonized inpatients. _Clostridium difficile_ levels in feces of infected patients are extremely high prior to treatment, decline during treatment, and rebound after the end of antibiotic treatment. High fecal bacterial counts are present for approximately 6 weeks after the end of clinical symptoms (26). A strong dose-response association between ward-level cohabitation with patients having symptomatic CDI and individual risk of future CDI has been shown (27,28). Furthermore, the number of hospital roommates (whether infected or not) has been shown to be a risk factor in one study (29), though recent research has shown that only a minority of CDI cases can be genetically linked to previously identified symptomatic cases (30).

Asymptomatic carriage in patients taking antibiotics is frequent, and such patients may pose an important but unidentified transmission risk in hospital (31).

Transmission events may occur via direct patient contact, via an initial contamination of a hospital surface, via the hands of health care workers, or via an airborne pathway, or some combination of the above (Figure 1.1). Due to the resilience of _Clostridium difficile_ spores, hospital surfaces, including floors and bedrails, harbor viable spores for extended periods after initial contamination. Research has shown that hospital surfaces are frequently contaminated, supporting the potential role of surfaces (32).
*Clostridium difficile* spores can be sampled from the air surrounding the majority of symptomatic CDI patients, and CDI spores remain suspended in the air for over 90 minutes after flushing a lidless toilet (33,34).

Studies of *Clostridium difficile* using environmental sampling techniques have not yet definitively linked CDI in humans with environmental spore density. Instead these studies have found that: (1) *Clostridium difficile* spores contaminate the direct surroundings in which patients are housed and the equipment used to care for them, (2) relatively less contamination occurs outside of CDI cases’ rooms, and (3) surfaces and equipment may act as reservoirs for future contamination (35).

On the other hand, an experimental animal study has shown that environmental contamination with *Clostridium difficile* spores was associated with CDI in a dose-dependent manner. Specifically, 4 spores per cm² were sufficient to infect 50% of clindamycin-prepared mice (2). To date, no studies have attempted to demonstrate a dose-response association between environmental spore exposure and CDI incidence using observational data of human infections. This may be due to the fact that studies have not developed methods for reliably assessing density of *Clostridium difficile* spores at relevant spatial and temporal scales.
1.1.5 Quantifying Risk of Transmission: Disease Pressure

Risk of cross infection for nosocomial pathogens can be incorporated into patient-level clinical prediction rules by measuring the magnitude and duration of exposure to infected or colonized patients. This effective measure is often called disease or colonization pressure (28) and is calculated as the number or proportion of patients on a given at-risk patients’ ward with the disease of interest. Past studies have considered disease pressure either as a risk factor in logistic regression models or as a time-varying covariate in Cox proportional hazards regression (28). As an example of the logistic regression strategy, *Clostridium difficile* disease pressure for a given patient was measured as the total number of patient-days of exposure to newly diagnosed CDI patients on a given ward (27). This study adjusted for individual level risk factors including patient age and antibiotic utilization. Patients with 2-8 patient-days’ exposure had a 3.9-fold (95% CI 2.8-5.5) increase in risk and patients with 9 or more patient-days’ exposure had a 9.7-fold (95% CI
7.1-13.1) increase in risk as compared to patients with 0 or 1 patient-days’ exposure to CDI infected patients. Disease pressure was most predictive of CDI risk, and more predictive than antibiotic exposure. Further, they found that the observed statistically significant effect of length of hospital stay was highly confounded by disease pressure: when both disease pressure and length of stay were included in a multivariate model, the effect of length of stay was not significant but disease pressure remained significant.

As an example of the time-varying covariate strategy, authors attempting to find risk factors for colonization with VRE in an ICU developed a measure of colonization pressure, equal to the proportion of patients with VRE (36) in the ICU on any given day. Each additional 1% increase in colonization pressure on a given day increased the hazards of VRE colonization by 3.2% (p=0.002), substantially more than enteral feeding and cephalosporin use.

1.1.6 Antibiotic-Associated Risk and the Role of Gut Flora
Antibiotic receipt represents the most important known risk factor for CDI. Antibiotic use is thought to induce CDI risk by denuding the gut of protective bacteria (37,38). However, what constitutes a healthy gut microbial ecosystem that is capable of resisting CDI is still poorly understood. Most evidence regarding the role of bacteria comes from experimental studies on the effectiveness of fecal transplant. Fecal transplant, involving implantation into the colon of feces obtained from a healthy donor, has been shown to be over three times more effective for curing patients with recurrent *Clostridium difficile* infection as compared to oral vancomycin therapy (6). Little is known about appropriate criteria for donor selection, but generally donors are family members of patients that have no recent history of antibiotic exposure. More recently, laboratory-produced combinations of
bacteria have been used to create synthetic gut microbial communities that have been shown to cure CDI (39). Nevertheless, little is known about what specific bacteria, or combinations of bacteria, confer resistance to CDI.

Recent research using 16S rRNA-encoding gene as a surrogate for individual bacterial species has attempted to correlate structural measures of gut microbial communities to resistance to *Clostridium difficile* colonization and virulence. One frequently cited article has shown that patients with recurrent CDI have lower microbiota diversity, characterized by an imbalance in Bacteroides and Firmicute species, though this study did not include any patients with a putatively healthy gut microbiology (i.e. there were no non-antibiotic controls) (40). Another recent study demonstrated that gut flora diversity (alpha-diversity based on rRNA) of patients with either fluoroquinolone or β-Lactam antibiotic exposures were similarly reduced relative to healthy controls that received no antibiotics (41).

Many empirical studies of antibiotic-associated risks in single hospitals have shown elevated risks of patients with recent antibiotic receipt relative to patients without this exposure. These studies have attempted to distinguish risk across specific antibiotic types, but are largely hindered by issues of statistical power (42). Further, antibiotic associated risks may depend on ward- or hospital-level antibiotic use since prevalent *Clostridium difficile* strains, and their resistance profiles may be impacted by antibiotic selection pressure (43). Larger scale studies of community-associated CDI across entire health care systems may provide better estimates of antibiotic associated risks due to the larger sample sizes and the lack of hospital-specificity of estimated effects. Relatively fewer studies have considered the duration of antibiotic associated risks.
antibiotic-specific risks and the duration of antibiotic risks are further discussed in the sections for Objective 1 and 2 of this thesis (below).

1.1.7 *Clostridium difficile* Seasonality

CDI incidence at the scale of hospital networks has a distinct annual seasonality. In Quebec, CDI incidence reaches a maximum incidence in the late spring of each year. This highly seasonal pattern was persistent despite a major outbreak from 2003 to 2006 (see Figure 1.2). The amplitude of CDI oscillation is substantial: peak incidence is approximately 50% higher than trough incidence.

Figure 1.2: *Clostridium difficile* infection seasonality in Quebec(21).

The seasonality of *Clostridium difficile* may be due to variation in antibiotic prescription associated with seasonal respiratory infections, particularly pneumonia and influenza. Pneumonia and influenza has been shown to be associated with CDI case counts in the United States, though this could potentially be due to increased numbers of hospitalizations in the winter (44). Other factors impacted by seasonal influenza epidemics, such as increased patient age in the wintertime and increased hospital occupancy (45), may also explain seasonality through mechanisms other than antibiotic
prescribing. Many upstream factors impacting the magnitude of wintertime pneumonia and influenza surges, such as influenza strain predominance (46), may also impact the magnitude of annual CDI surges.

1.1.8 Other CDI Risk Factors

Notwithstanding inter-patient transmission and antibiotic-associated risk, other risk factors for CDI resemble those for other nosocomial pathogens, and include age, recent history of hospitalization, and underlying disease, including renal failure and severity of chronic illness which can be measured by the Charlson comorbidity index (47,48). Invasive procedures or devices including gastrointestinal surgery, transplantation, central venous or arterial catheter, urinary catheter, intubation and mechanical ventilation, and tube feeding (49) may increase risk by increasing the quantity of viable spores in the digestive tract.

Various pharmacologic exposures including immune suppressant use, laxative use and gastric acid suppressant use are also thought to increase risk. Gastric acid suppressant use is thought to increase risk due to higher survival of ingested Clostridium difficile spores, though evidence is conflicting (50).

Dubberke et al. (51) developed a clinical prediction model based on eight risk factors: Clostridium difficile disease pressure, age, the number of times admitted to hospital in the previous 60 days, modified Acute Physiology Score, days of treatment with high-risk antibiotics, albumin level, admission to an intensive care unit, and receipt of laxatives, gastric acid suppressors, or antimotility drugs. However, their model yielded a low level of statistical discrimination (c=0.71), indicating low sensitivity and specificity for
predicting infection. As such, there remains substantial room for improving our knowledge of *Clostridium difficile* risk factors and developing new predictive models.

1.1.9 Statistical Analysis of Hospital-Acquired Infections

Several statisticians have brought forward objections to current statistical practices with respect to the evaluation of risk factors for hospital-acquired infections and the effectiveness of infection control interventions. The specific concerns that have been raised are related to 1) time-dependent and length bias (52), 2) biases due to dependent censoring (52), and 3) lack of control for correlated outcomes (53). Time-dependent bias and length bias occur in hospital-based studies when a given risk factor that is transient (such as an antibiotic exposure) or not present on admission (such as a hospital acquired infection) is assumed to have acted over the course of the entire hospitalization period. Dependent censoring is often present in studies of hospital-acquired infections as patients discharged tend to be healthier than patients that remain hospitalized for longer periods. Publications on the effects of a given infection control measure in the infection control literature are often based on short time series of low-numbered counts of infected patients (often from single institutions) and most analytic methods ignore the correlated nature of the count observations (53). A systematic review of evidence for MRSA isolation policies found that of 24 studies presenting statistical analyses, all but one assumed outcomes to be independent (54). Analytic methods used to understand the evolution of infectious disease rates and to establish the impact of interventions could be improved by using previously developed methods that take these issues into account, including the use of time-dependent exposure variables, accounting for correlated outcomes using GEE or multilevel modeling approaches, and inverse-probability of censoring weights.
**Distributed Lag Models and Weighted Cumulative Exposures**

Distributed lag models were developed by econometricians in order to measure the lag-specific effect size of a time-varying exposure on an outcome. Almon (55) studied the effects of capital appropriations on expenditures, and proposed a specific distributed lag structure wherein weights of lagged capital appropriations would follow a polynomial curve with a pre-specified maximum lag. For example, the best distributed lag model for the effect of capital appropriations on expenditures in the manufacturing industry used a pre-specified maximum lag of 7 months and estimated a peak at 3 months lag, with earlier and later effects being smaller (Figure 1.3). Distributed lag non-linear models (DLNM) are an extension to the Almon model, which allow for non-linear effects and a wider range of distributed lag effects (56). These models are fitted by transforming the lag covariates into a linearly independent spanning set, known as a cross-basis. The cross-basis is then added to the linear model as a set of covariates. The statistical significance of a given set of lagged effects is measured by a likelihood ratio test (LRT) with the same number of degrees of freedom as the size of the spanning set. Such distributed lag non-linear models have been used in time-series models to assess the impact of time-specific exposures on health outcomes such as the impact of heat waves on mortality (Figure 1.4) (57). The methods associated with DLNMs can also be used in the context of individual-level Cox or logistic regression models: in these cases, they are usually called weighted cumulative exposures (58). In the context of health-related research, these methods have been applied by pharmacoepidemiologists to determine the delayed impacts of benzodiazepines on fall-related injuries (59) and by environmental health researchers to consider the delayed impacts of heat waves on all-cause mortality (57).
1.2  Note on the Format of the Thesis

This dissertation is presented in the journal-based format in which the material is presented as a series of stand-alone manuscripts. Each of the specific thesis objectives (below) corresponds to a specific manuscript. A reader’s note introduces each manuscript while supplementary methods and results are presented after the discussion section.

1.3  Thesis Objectives and their Scientific Contribution

The global purpose of this thesis is to improve the understanding of the determinants of time-trends in *Clostridium difficile* incidence at a population-level and to better our understanding of the role of antibiotics in the occurrence of CDI through improved statistical modeling of the timing of CDI acquisition.
1.3.1 Objective 1: Systematic Review on Antibiotics and Community-Associated CDI

Several systematic reviews have considered the association between antibiotic exposures and the risk of developing CDI (42,60). These studies have shown that differential risks may exist between antibiotic classes. On the basis of risk of CDI in patients, certain antibiotic classes have been considered as high-risk (most commonly: clindamycin, 3rd generation cephalosporins, fluoroquinolones), though studies differ in their classifications (51,61). In most hospital-based studies, certain infrequently prescribed antibiotics are likely misclassified as either high- or low-risk due to lack of statistical power. Studies of community-associated CDI may be more reliable, as they use state-level data that spans many hospital outpatient groups and regions. Thus, a systematic review of community-associated CDI risk could provide the most reliable evidence and classification of antibiotic-associated CDI risk. I propose to conduct a systematic review of the association between antibiotic use and the risk of community-associated CDI. A reliable classification of antibiotic risk for CDI would enable clinicians to restrict high-risk antibiotics during community CDI outbreaks and improve patient outcomes.

Table 1.1: Thesis objectives and research questions

<table>
<thead>
<tr>
<th>Objective</th>
<th>Research questions</th>
<th>Chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. To quantify the effects of antibiotics on risk of CDI among non-hospitalized patients</td>
<td>1.1. What are the risks associated with exposures to the following 7 classes of antibiotics: (i) tetracyclines; (ii) sulfonamides and trimethoprim; (iii) penicillins; (iv) macrolides; (v) cephalosporins, monobactams and carbapenems (CMCs); (vi) fluoroquinolones; and (vii) clindamycin</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1.2. Can a simple index be used to summarize the risk observed between antibiotic classes?</td>
<td></td>
</tr>
<tr>
<td>2. To assess the degree to</td>
<td>2.1. How long does risk last after the end</td>
<td>3</td>
</tr>
</tbody>
</table>
which risks associated with antimicrobial exposures both cumulate over the course of antimicrobial therapy and abate after cessation of antibiotic exposure?

2.2. Does risk increase over the course of antibiotic therapy?

3. To evaluate the link between the seasonality of P&I and hospital-associated CDI at a national level

3.1 What is the mean annual timing of the CDI peak and the mean annual timing of the P&I peak, and what is the time elapsed between the two?

3.2 What is the impact of an increase in hospital P&I prevalence on subsequent CDI incidence?

1.3.2 Objective 2: Duration of CDI Risk Associated with Antibiotic Exposure

Several studies have followed patients for periods of 30 to 100 days post-admission and shown that a relatively elevated incidence of CDI exists for patients post-discharge (62–65). Further, recent studies have used survival analysis incorporating time-varying antibiotic exposures and indicate that increased duration, number, and dosages of antibiotics are associated with increased risk (61,66). However, few studies have attempted to identify the magnitude of risk throughout (before, during, and after) antimicrobial exposure. Weighted cumulative exposure models build on survival analysis and stipulate that current risks may be considered as functions of past exposures (59), and may be used to explicitly and flexibly estimate the day-by-day CDI risk during and after antibiotic therapy. As such, my objective was to assess the degree to which risks associated with antimicrobial exposures both accumulate over the course of antimicrobial therapy and abate after cessation.
1.3.3 Objective 3: Population-Level Seasonal CDI Variation

Recently, CDI incidence has been shown to be associated with the seasonal variations in the incidence of pneumonia and influenza hospitalizations (44). Increased wintertime prescription of broad-spectrum antibiotics has been suggested as a causal mechanism explaining this co-seasonality. Studies of seasonal patterns of influenza and pneumonia hospital admissions demonstrate large wintertime increases in pneumonia and influenza related hospitalizations, particularly among infants and the elderly (45). Influenza-related admissions are increasingly concentrated in the elderly, and this has been attributed to the increasing frequency of A(H3N2) strain-dominant seasons and the decreasing severity of influenza A(H3N2) seasons for the younger age-groups (46). Studies developing methods for linking wintertime increases in pneumonia and influenza (P&I (67)) hospitalizations to circulating viruses have been critiqued since they are generally based on time-series analyses with time-lagged measures of circulating influenza A, B and RSV, and have not been validated with individual-level viral testing (68). In this study, my objective was to consider the link between CDI incidence and the prevalence of P & I in a representative sample of US hospitalizations over a 16-year period.

1.4 References


52. Schumacher M, Allignol A, Beyersmann J, Binder N, Wolkewitz M. Hospital-acquired infections--appropriate statistical treatment is urgently needed! Int J Epidemiol. 2013 Sep 14;


Chapter 2. Antibiotics and the risk of community-associated Clostridium difficile infection (CDI): a meta-analysis

2.1 Reader’s Note

Many studies have considered the variation in risk associated with different classes of antibiotics; but many of these are plagued by small sample sizes. Here I present evidence that certain classes of antibiotics, those that also happen to be the most frequently prescribed in hospitals, are associated with the highest risk. A version of this paper was published in the journal Antimicrobial Agents and Chemotherapy.


2.2 Abstract

The rising incidence of Clostridium difficile infection (CDI) could be reduced by lowering exposures to high risk antibiotics. The objective of this study was to determine the association between antibiotic class and the risk of CDI in the community setting. EMBASE and PubMed were queried without restriction to date or language. Comparative observational studies and RCTs considering the impact of antibiotic exposures on CDI risk among non-hospitalized populations were considered. We estimated pooled odds ratios (OR) for antibiotic classes using random effects meta-analysis. Our search criteria identified 465 articles, of which 7 met inclusion criteria; all were observational studies. Five studies considered antibiotic risk relative to no antibiotic exposure: clindamycin (OR=16.80, 95% confidence interval [CI]: 7.48–37.76), fluoroquinolones (OR=5.50, 95% CI: 4.26–7.11) and cephalosporins, monobactams and carbapenems (CMCs, OR=5.68, 95%
CI: 2.12–15.23) had the largest effects, while macrolides (OR=2.65, 95% CI: 1.92–3.64), sulfonamides and trimethoprim (OR=1.81, 95% CI: 1.34–2.43) and penicillins (OR=2.71, 95% CI: 1.75–4.21) had lesser associations with CDI. We noted no effect of tetracyclines on CDI risk (OR=0.92, 95% CI: 0.61–1.40). In the community setting, there is substantial variation in risk of CDI associated with different antimicrobial classes. Avoidance of high risk antibiotics (such as clindamycin, CMCs and fluoroquinolones) in favor of lower risk antibiotics (such as penicillins, macrolides and tetracyclines) may help reduce the incidence of CDI.

2.3 Introduction

Clostridium difficile, a toxin-producing bacterium that causes diarrhea, is the largest single cause of morbidity and mortality among hospital-acquired infections (1). In-hospital, C. difficile infection (CDI) is generally acquired when patients with predisposing factors such as advanced age and antibiotic use are exposed to C. difficile spores emanating from other hospitalized infected patients (2). With the emergence of increasingly virulent C. difficile strains have come reports of CDI among patients previously considered to be at low risk of this infection, including those living in the community (3–5). Spore exposure may occur outside inpatient settings since river water, soil and foods can be contaminated (6, 7), outpatient exposures to the healthcare system are common, and transmission may occur within households (8). A recent study noted that the population-based incidence of community-acquired CDI (11.2 cases per 100,000 person-years) was on par with hospital-acquired CDI (12.1 cases per 100,000 person-years) (9).

One published meta-analysis and one systematic review have considered the risk of antibiotic exposures for CDI (10, 11) risk among hospital inpatients. The meta-analytic
study noted that tetracyclines and penicillins were associated with the lowest risk, while fluoroquinolones, clindamycin and 3rd generation cephalosporins were associated with the highest risk of CDI acquisition, despite considerable confidence interval overlap (10). The systematic review established that the strongest evidence of risk existed for penicillins and clindamycin and that effect estimates for other antibiotic classes were liable to bias (11).

In addition to yielding accurate adjusted effect estimates, a systematic review of the association between antibiotic exposures and community-associated CDI is necessary since the risk profile is different among non-hospitalized populations (younger age, less frequent exposure to patients with symptomatic CDI, and different profile of underlying infections and antibiotic treatments). We conducted a systematic review of the association between antibiotic type and the risk of CDI in non-hospitalized populations. Our objective was to quantify the relative risks of specific antibiotics in order to better understand the risks of prescribing various antibiotics in the community setting.

2.4 Methods

2.4.1 Search Criteria

A literature search was conducted in March 2012 using the EMBASE and PubMed databases and included all articles without restriction to language or time period. References of included articles were browsed and content experts were approached to identify further relevant articles. Within each database, our search strategy was to use both keywords and mapped subject headings as terms describing the exposure (i.e. antibiotic, antibacterial, antimicrobial, aminoglycosides, beta-lactams, cephalosporins, clindamycin, fluoroquinolones, macrolides, metronidazole, sulfonamides, and
tetracyclines), outcome (\textit{C. difficile} infection) and the detection of a community-acquired infection (community-acquired, community-associated, outpatient, ambulatory care, registry, and general practice). Exposure, outcome and population terms were then combined using the Boolean “AND” operator (12).

We included population-based studies of people with little to no health care system contact prior to the onset of disease (13); studies restricted to hospital acquired or healthcare associated disease, (e.g. studies restricted to HIV and cancer outpatients) were excluded. The outcome of interest was incident CDI (collected at the individual level). Exposures of interest were specific antibiotic classes. We excluded studies considering risk factors for severe or relapsing CDI among individuals already diagnosed with \textit{C. difficile}, time-series analyses, and those not examining specific antibiotics or antibiotic classes.

### 2.4.2 Screening and Data Abstraction

One author (KAB) screened article titles and abstracts of the initial database search so as to identify those appropriate for full text review. Full text of identified articles was read; those articles eliminated in the initial screen of titles and abstracts were distinguished from articles eliminated in the full-text screen. Data on numbers of cases and controls; unadjusted and adjusted effect sizes; and 95% confidence intervals corresponding to each antibiotic exposure group reported, were abstracted and entered into a spreadsheet. When insufficient information was available to obtain the appropriate effect size standard errors, study authors were contacted by email.
2.4.3 Quality Assessment

Study quality was assessed using a 6 criterion checklist (14–16) with certain elements of the checklist aimed to address specific concerns raised in previous systematic reviews on the topic (Table 2.1) (10, 11). Each identified quality criterion except question 3 was scored as 0 (no) or 1 (yes); question 3 was graded on a zero to 2 scale. Two authors (KAB and NK) independently graded study quality; the results were compared and disagreements were resolved through discussion. By summing up the points, the studies were classified as high quality (6-7 points), moderate quality (4-5 points) or poor quality (< 4 points).

Table 2.1. Study quality assessment questions.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Did the study have a clearly stated study aim?</td>
</tr>
<tr>
<td>2.</td>
<td>Was the study population clearly defined (i.e. was an appropriate method used to ensure that inpatients were excluded)?</td>
</tr>
<tr>
<td>3.</td>
<td>Were the antibiotic exposure groups case counts reported, and if so, were they sufficiently large to statistically compare effect sizes (2: ≥10 cases in each antibiotic exposure group; 1:&lt;10 cases in some; 0: not reported)?</td>
</tr>
<tr>
<td>4.</td>
<td>Were the criteria for diagnosis of <em>C. difficile</em> clearly defined (e.g., statement of ICD code for registry studies; clear description of microbiologic methods used for clinical studies)?</td>
</tr>
<tr>
<td>5.</td>
<td>Was there an attempt to use statistical analyses to adjust for confounding or to standardize disease rates by age, duration of antibiotic exposure and exposure to combinations of antibiotics?</td>
</tr>
<tr>
<td>6.</td>
<td>Was there a discussion of study limitations?</td>
</tr>
</tbody>
</table>
2.4.4 Variables

Antibiotics were classified into one of the following 7 groups as per included articles: (i)tetracyclines; (ii)sulfonamides and trimethoprim; (iii)penicillins; (iv)macrolides; (v)cephalosporins, monobactams and carbapenems (CMCs); (vi)fluoroquinolones; and (vii)clindamycin. The dependent variable of interest was the adjusted study-specific log odds-ratio of a given antibiotic class relative to no antibiotic exposure; this variable and the standard error for each reported effect were usually transcribed directly from the publication results section. In Dial et al. (17), effect sizes of levofloxacin, gatifloxacin and moxifloxacin antibiotics were combined by taking the weighted average of the log-odds ratios, with inverse variance weights; in Kuntz et al. (9), the effect sizes for the CMCs were combined in a similar manner.

2.4.5 Statistical Analysis

A pooled random-effects analysis was used to consider the impact of any antibiotic exposures irrespective of antibiotic class relative to no exposure, using the Dersimonian-Laird approach (18). A stratified analysis was then used to consider the risk associated with each antibiotic type compared to no antibiotic exposure. In a secondary analysis restricted to the 2 studies excluding antibiotic-unexposed persons (19, 20), odds ratios were calculated from the number of CDI-positive cases and the total person-time within antibiotic exposure classes with 0.5 added to each cell (in order to estimate effects in the presence of zero cells) (21); for the odds ratio, penicillin-exposed persons were considered as the referent category. Due to the low incidence of community-acquired CDI in contemporary populations (9), the odds ratios reported in this study may be interpreted
as risk ratios (22). For these studies, a similar stratified meta-analysis measured the risk associated with each antibiotic type compared to penicillin exposure.

We assessed the heterogeneity of study results by use of the $\tau^2$, $Q$ (18), and $I^2$ statistics (23). Possible sources of heterogeneity were explored in sensitivity analyses in which certain subgroups were excluded, as well as through creation of meta-regression models (24). Analyses were conducted in R using the meta and metafor packages (25).

2.5 Results
Out of a total of 465 articles identified, only 7 fulfilled the eligibility criteria (Fig. 1). Three studies employed nested case-control designs, 3 used non-nested case-control designs and one was a cohort study (Table 2.2). The studies followed subjects between 1988 and 2007, and varied in size from as few as 40 to over 1200 cases of CDI.
Figure 2.1: Flow chart of studies screened and included.

Potentially relevant studies identified and screened for retrieval (n=465)

Studies excluded (n=435):
• not original research article (220)
• nosocomial or healthcare associated infections (75)
• restricted to cases of CDI (61)
• not CDI outcome (59)
• restricted to children (12)
• no specific antibiotic exposure reported (4)
• other (4)

Studies retrieved for full text review (n=30)

Studies included in the systematic review (n=7)

Studies excluded (n=23):
• no specific antibiotic classes (8)
• case only (5)
• not community acquired cases (3)
• repeat analysis on same dataset, time series, not research article, not CDI, restricted to fluoroquinolone exposures, cancer outpatients, matched on antibiotic class exposure (1 each)

Studies included in primary meta-analysis (n=5)

Studies without antibiotic free referent (n=2)
• included in secondary meta analysis (2)
<table>
<thead>
<tr>
<th>Study Years</th>
<th>Case definition</th>
<th>Matching</th>
<th>Adjustment method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 1992 - Dec 1994</td>
<td>Presence of a first positive <em>C. difficile</em> toxin assay and/or a clinical diagnosis recorded by their general practitioner</td>
<td>None</td>
<td>Conditional logistic regression</td>
</tr>
<tr>
<td>Apr 1988 - Nov 1990</td>
<td>Positive assay for <em>C. difficile</em> toxin with documented diarrhea or colitis with onset within 48h of admission</td>
<td>None</td>
<td>Conditional logistic regression</td>
</tr>
<tr>
<td>Jan 2004 - Dec 2007</td>
<td>Primary or secondary CDI diagnosis on insurance claim. For hospital patients, must be recorded at admission</td>
<td>None</td>
<td>Conditional logistic regression</td>
</tr>
<tr>
<td>Jan 1996 - Dec 2004</td>
<td>Documented diarrhea and a positive <em>C. difficile</em> laboratory test result in the medical record</td>
<td>Index date** and date of 1st hospitalization</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>Jan 1993 - Dec 2004</td>
<td>Patients with diarrhea (onset in community or ≤ 72 hours of hospital arrival) and a positive stool toxin assay</td>
<td>None</td>
<td>Conditional logistic regression</td>
</tr>
<tr>
<td>Jan 1999 - Dec 2007</td>
<td>Patients presenting to GP with symptoms of diarrhea with positive stool toxin assay</td>
<td>Index month**, clinic site</td>
<td>Matching</td>
</tr>
</tbody>
</table>

**Matching:** Age, clinic site and index date**

**Adjustment method:**
- Conditional logistic regression
- Matching

<table>
<thead>
<tr>
<th>Date</th>
<th>Source</th>
<th>Cases (n) / Controls (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2/3/12 12/30/00</td>
<td>1,233 / 12,330</td>
<td></td>
</tr>
<tr>
<td>8/6/30/00</td>
<td>836 / 8,360</td>
<td></td>
</tr>
<tr>
<td>5/17/37/00</td>
<td>51 / 175,275</td>
<td></td>
</tr>
<tr>
<td>8/6/30/00</td>
<td>304 / 3,040</td>
<td></td>
</tr>
<tr>
<td>5/17/37/00</td>
<td>48 / 358,389</td>
<td></td>
</tr>
<tr>
<td>1/2/3/12 12/30/00</td>
<td>66 / 114</td>
<td></td>
</tr>
</tbody>
</table>

<p>|</p>
<table>
<thead>
<tr>
<th>Date</th>
<th>Source</th>
<th>Cases (n) / Controls (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/2/00 - Nov 2007</td>
<td>112 / 1,112</td>
<td></td>
</tr>
<tr>
<td>9/8/36/98</td>
<td>48 / 3,389</td>
<td></td>
</tr>
<tr>
<td>5/15/37/00</td>
<td>34 / 1,374</td>
<td></td>
</tr>
<tr>
<td>8/6/30/00</td>
<td>24 / 2,140</td>
<td></td>
</tr>
<tr>
<td>5/17/37/00</td>
<td>11 / 7,392</td>
<td></td>
</tr>
<tr>
<td>8/6/30/00</td>
<td>9 / 1,911</td>
<td></td>
</tr>
</tbody>
</table>

* VAMC: Veterans Affairs Medical Center, CIHI: Canadian Institute for Health Information.
A total of 5 studies included controls without antibiotic-exposure and could be included in the primary meta-analyses; the other 2 studies were included in a secondary meta-analysis limited to studies without antibiotic free controls.

Of the studies in the primary analysis, two evaluated all 7 candidate antibiotic classes, two covered between 5 and 6 of the 7, and one reported on only 4 of 7. The two studies in the secondary analysis each reported exposures for all 7 antibiotic classes, but no patients exposed to clindamycin in either of these additional studies acquired CDI infection and so odds ratios were not calculated for this agent.

Among the studies included in the primary meta-analyses, study quality (Table 2.3) was scored high for 2 studies, moderate for 2, and low for one. Among studies included in the secondary meta-analyses, one was scored as high quality, and the other was scored as low quality. Note that for three studies, the case definition did not properly exclude potentially hospital-acquired cases (17, 26, 27).
Table 2.3: Quality characteristics of eligible studies.

<table>
<thead>
<tr>
<th>First Author (Publication Year)</th>
<th>Quality Score (0-7)</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3*</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delaney (2007)</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dial (2008)</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hirschhorn (1994)</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Kuntz (2011)</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Levy (2000)</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Naggie (2011)</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Wilcox (2008)</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

* Scored on a 0-2 scale.

2.5.1 Pooled Effects

The pooled impact of any antibiotic exposure (irrespective of antibiotic class) across all 29 antibiotic effects (Fig. 2) was to increase risk of CDI by a multiple of 3 (OR=3.55, 95% CI: 2.56—4.94). In this analysis, which ignored antibiotic class, effect heterogeneity was extremely high ($I^2=90.6\%$, p<0.001).
Figure 2.2: Pooled and study-specific risk estimates of community-associated CDI risk by antibiotic class.

<table>
<thead>
<tr>
<th>First Author (Year)</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tetracyclines</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delaney (2007)</td>
<td>0.90</td>
<td>(0.54-1.50)</td>
</tr>
<tr>
<td>Dial (2008)</td>
<td>1.10</td>
<td>(0.14-8.60)</td>
</tr>
<tr>
<td>Kuntz (2011)</td>
<td>0.94</td>
<td>(0.43-2.04)</td>
</tr>
<tr>
<td><strong>Combined odds ratio</strong></td>
<td>0.92</td>
<td>(0.61-1.40)</td>
</tr>
<tr>
<td><strong>Penicillins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delaney (2007)</td>
<td>1.90</td>
<td>(1.50-2.40)</td>
</tr>
<tr>
<td>Dial (2008)</td>
<td>4.30</td>
<td>(2.89-6.40)</td>
</tr>
<tr>
<td>Kuntz (2011)</td>
<td>1.72</td>
<td>(1.16-2.54)</td>
</tr>
<tr>
<td>Naggie (2011)</td>
<td>3.38</td>
<td>(1.55-7.37)</td>
</tr>
<tr>
<td>Wilcox (2008)</td>
<td>6.50</td>
<td>(1.60-26.48)</td>
</tr>
<tr>
<td><strong>Combined odds ratio</strong></td>
<td>2.71</td>
<td>(1.75-4.21)</td>
</tr>
<tr>
<td><strong>Sulfonamides and trimethoprim</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delaney (2007)</td>
<td>1.90</td>
<td>(1.34-2.70)</td>
</tr>
<tr>
<td>Dial (2008)</td>
<td>1.20</td>
<td>(0.44-3.30)</td>
</tr>
<tr>
<td>Kuntz (2011)</td>
<td>1.58</td>
<td>(0.79-3.15)</td>
</tr>
<tr>
<td>Wilcox (2008)</td>
<td>5.45</td>
<td>(0.75-39.86)</td>
</tr>
<tr>
<td><strong>Combined odds ratio</strong></td>
<td>1.81</td>
<td>(1.34-2.43)</td>
</tr>
<tr>
<td><strong>Macrolides</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delaney (2007)</td>
<td>2.20</td>
<td>(1.56-3.10)</td>
</tr>
<tr>
<td>Dial (2008)</td>
<td>3.90</td>
<td>(2.58-5.90)</td>
</tr>
<tr>
<td>Kuntz (2011)</td>
<td>2.19</td>
<td>(1.54-3.11)</td>
</tr>
<tr>
<td>Wilcox (2008)</td>
<td>4.01</td>
<td>(0.79-20.48)</td>
</tr>
<tr>
<td><strong>Combined odds ratio</strong></td>
<td>2.65</td>
<td>(1.92-3.64)</td>
</tr>
<tr>
<td><strong>Cephalosporins, monobactams and carbapenems (CMCs)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delaney (2007)</td>
<td>2.20</td>
<td>(1.51-3.20)</td>
</tr>
<tr>
<td>Dial (2008)</td>
<td>14.90</td>
<td>(10.94-20.30)</td>
</tr>
<tr>
<td>Naggie (2011)</td>
<td>8.84</td>
<td>(1.85-42.30)</td>
</tr>
<tr>
<td>Wilcox (2008)</td>
<td>6.49</td>
<td>(1.42-29.73)</td>
</tr>
<tr>
<td>Kuntz (2011)</td>
<td>3.77</td>
<td>(2.35-6.04)</td>
</tr>
<tr>
<td><strong>Combined odds ratio</strong></td>
<td>5.68</td>
<td>(2.12-15.23)</td>
</tr>
<tr>
<td><strong>Fluoroquinolones</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dial (2008)</td>
<td>6.05</td>
<td>(3.68-9.94)</td>
</tr>
<tr>
<td>Kuntz (2011)</td>
<td>4.91</td>
<td>(3.28-7.35)</td>
</tr>
<tr>
<td>Naggie (2011)</td>
<td>1.31</td>
<td>(0.28-6.04)</td>
</tr>
<tr>
<td>Wilcox (2008)</td>
<td>9.39</td>
<td>(0.98-90.05)</td>
</tr>
<tr>
<td><strong>Combined odds ratio</strong></td>
<td>5.50</td>
<td>(4.26-7.11)</td>
</tr>
<tr>
<td><strong>Clindamycin</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dial (2008)</td>
<td>31.80</td>
<td>(17.56-57.60)</td>
</tr>
<tr>
<td>Kuntz (2011)</td>
<td>13.00</td>
<td>(7.03-24.04)</td>
</tr>
<tr>
<td>Naggie (2011)</td>
<td>6.84</td>
<td>(1.34-33.00)</td>
</tr>
<tr>
<td><strong>Combined odds ratio</strong></td>
<td>16.80</td>
<td>(7.40-37.70)</td>
</tr>
<tr>
<td><strong>Pooled odds ratio</strong></td>
<td>3.55</td>
<td>(2.56-4.94)</td>
</tr>
</tbody>
</table>
2.5.2 Antibiotic Types

In the analyses stratified by antibiotic class, 6 of 7 antibiotic classes were associated with increased risk of CDI (Fig. 2). Specifically, clindamycin (OR=16.80, 95% CI: 7.48–37.76), fluoroquinolones (OR=5.50, 95% CI: 4.26–7.11) and CMCs (OR=5.68, 95% CI: 2.12–15.23) were found to increase CDI risk the most, while macrolides (OR=2.65, 95% CI: 1.92–3.64), sulfonamides and trimethoprim (OR=1.81, 95% CI: 1.34–2.43) and penicillins (OR=2.71, 95% CI: 1.75–4.21) were found to have a lesser, but nevertheless statistically significant impact. There was no evidence of the impact of tetracyclines on CDI risk (OR=0.92, 95% CI: 0.61–1.40).

Between-study heterogeneity was largest in the CMCs ($I^2=93.8\%$, $p<0.001$), penicillins ($I^2=76.8\%$, $p=0.002$) and clindamycin ($I^2=66.7\%$, $p=0.05$) drug classes. Conversely, heterogeneity was lowest for tetracyclines ($I^2=0.0\%$, $p=0.98$), sulfonamides and trimethoprim ($I^2=0.0\%$, $p=0.56$) and fluoroquinolones ($I^2=10.9\%$, $p=0.34$). Relative to the pooled meta-analysis ($\tau^2 = 0.62$), the stratified model reduced heterogeneity by 55% ($\tau^2 = 0.27$, $p < 0.001$).

2.5.3 Meta-Regression

Meta-regression was used in order to investigate the factors influencing residual heterogeneity from the primary analysis. The five studies were associated with systematic differences in drug effects ($\chi^2_{4\text{df}}, p = 0.001$); in particular, the pooled odds ratios in one study(17) were twice those of the remaining studies (OR= 1.93, 95% CI: 1.30—2.87). The addition of study level effects to the meta-regression model reduced heterogeneity by 63% ($\tau^2 = 0.10$, $p < 0.001$). As a sensitivity analysis, we excluded the one study reporting larger
effect sizes; the pooled odds ratio in the remaining 4 studies was 2.86 (95% CI: 2.86–3.81) and the between study effects were no longer significant ($\chi^2_{3df}, p = 0.12$).

For the subset of high quality studies (n=2), heterogeneity of between study effects was below detectable limits ($\chi^2_{1df}, p = 0.96$). These two studies reported effect sizes for all antibiotic classes that were smaller, relative to those of medium and low quality studies (OR=2.50, 95% CI 1.80–3.47).

### 2.5.3 Publication Bias

We tested for funnel plot asymmetry using the stratified model and found no evidence of an association between effect estimate precision and residual effect sizes ($z = 0.53, p = 0.59$).

### 2.5.4 Antibiotic-Associated CDI Risk (AACR) Index

In a post-hoc exploratory analysis, a simple 4 point index summarizing the meta-analysis results was developed; the index was equal to 1 for tetracyclines, 2 for sulfonamides and trimethoprim, macrolides and penicillins, 3 for CMCs and fluoroquinolones and 4 for clindamycin. Each one point increase in the index was associated with a 2.41 (95% CI: 2.14—2.74) fold increase in the odds of acquiring CDI. Mean antibiotic class effect did not deviate significantly from the linear trend ($p_{sulfanomides}=0.30$ to $p_{tetracyclines}=0.85$). The model fit is presented graphically in Figure 3.
2.5.5 Secondary Analysis

With this analysis, similar findings were noted; namely, tetracyclines (OR=0.60, 95% CI: 0.14—2.61) were not associated with an increased risk of *C. difficile*; sulfonamides and trimethoprim (OR=0.85, 95% CI: 0.29—2.52) and macrolides (OR=0.60, 95% CI: 0.20—1.76) tended to have the smallest effect sizes with the least heterogeneity; while CMCs (OR=4.12, 95% CI: 2.28—7.44) and fluoroquinolones (OR=4.31, 95% CI: 1.46—12.70) had larger and more variable effect sizes. In both studies, clindamycin exposure was rare (<0.5% of total antibiotic exposures in each); neither reported any cases associated with clindamycin.

2.6 Discussion

The emergence of *C. difficile* as an infection in individuals without prior hospitalization, and presumed community acquisition, represents a concerning development in the ongoing emergence of this pathogen. As any prescription of an antimicrobial agent to a
patient in an outpatient setting requires a careful weighing of risks and benefits, we
performed a systematic review to quantify the risks associated with individual antibiotic
classes, and to identify areas of heterogeneity in such risk. Overall use of antibiotic agents
is associated with a threefold increased risk of community-acquired CDI, but we also
detected substantial variation in risk associated with different antimicrobial classes, with
fluoroquinolones, CMCs, and clindamycin associated with the greatest enhancement of
risk.

This study largely corroborates the associations found for hospital-associated CDI
risk (10, 11, 28). In keeping with many historic studies of CDI risk and outbreaks of the
disease, clindamycin was found to have the strongest association with risk. One must note
however, that clindamycin has not been associated with the greatest risk enhancement in
every study (28); variability in effects may be due to true biological heterogeneity of effect
(e.g. variable strain susceptibility to clindamycin (29), timing of inoculation relative to the
end of antibiotic exposure), or could be an artifact of the different methods used for
outcome ascertainment (see below).

Our study found large effects for fluoroquinolones and CMCs. This could be
expected given the broad spectrum of activity of these agents against intestinal microbes,
and low susceptibility of Clostridium difficile to these classes of antibiotics (30). The risk
associated with CMCs was highly variable across studies, in contrast to fluoroquinolones,
which appeared to have more consistent effects. This heterogeneity may be due to the
greater activity of later generation cephalosporins against anaerobic bacteria and gram
negative bacilli (10). In contrast, one study limited to patients with fluoroquinolone
exposures (31), found no differences in effect between levofloxacin, gatifloxacin and moxifloxacin, notwithstanding the enhanced anti-anaerobic spectrum of the latter agents.

Our findings reaffirmed the finding of moderate effects for penicillins, macrolides and sulfonamides and trimethoprim from a recent hospital-based study (28); this is in contrast to other hospital-based studies that have noted large effects for penicillins (11). The relatively low minimum inhibitory concentration (MIC) for penicillins among common CDI strains (32) could help explain the observed modestly elevated risk level for the penicillin class, such that the enhancement of CDI risk resulting from elimination of normal enteric flora is somewhat counterbalanced by anti-\textit{C. difficile} activity. These discrepancies may also result from wide variations in the antibiotic spectrum of penicillin subclasses (including broad-spectrum penicillins used more in the hospital setting such as piperacillin-tazobactam). Our meta-analysis noted that tetracycline antibiotics have little antibiotic-associated risk, which is in keeping with the only meta-analysis of inpatient CDI (10) risk.

Like any observational study, the findings of studies incorporated into this meta-analysis could have been biased by methodological flaws, including issues of control selection, misclassification of both outcomes and exposures and residual confounding (33). Our quality checklist attempted to assess the overall risk of these biases in each study; we outline some specific observations below. With respect to the definition of the population, two studies did not exclude patients exposed to hospital settings during the risk period, and as such may actually represent studies of community-onset but hospital-acquired disease (20, 27) while two studies were restricted to patients who received a \textit{C. difficile} assay, and as such controls did not represent the source population of cases (20,
With respect to the outcome ascertainment, all positive cases may have been subject to misclassification due to co-infection with another diarrhea-causing organism. Further, in two studies (17, 20), a lack of clinical detail meant that hospital-diagnosed cases with onset of symptoms ≥ 48h after admission could not be separated from those with onset within 48h. As such, unmeasured inpatient antibiotic exposures may have caused the disease outcome. Indeed, in Dial et al. (17) outpatient antibiotic exposures were only detected in 47.1% of cases. Of included studies, only one (34) considered the robustness of results to diagnostic suspicion bias by comparing effect sizes from clinical diagnoses to those with test-based confirmation; they found no significant differences in effect with the clinically diagnosed subgroup.

Other potential sources of bias in our meta-analysis could include a lack of consensus regarding the appropriate time-window for identification of antibiotic exposure. As risk associated with antibiotic exposure decreases with increasing time, larger effect sizes are liable to be found in studies looking at the shorter time-windows. Indeed, the study in our primary analyses with the shortest exposure window reported larger effect sizes for all antibiotics except tetracyclines (17). In fact, the appropriate time-window may differ between antibiotic classes, due to differing antimicrobial effect duration (35). In addition, simultaneous administration of multiple antimicrobial agents and confounding by indication (as individuals receiving antimicrobials may have underlying health conditions placing them at greater risk for CDI) may have biased results.

Finally, the relatively small number of studies that met the inclusion criteria was a weakness of this systematic review. And, although we did not find evidence to suggest that our findings were influenced by publication bias, we did notice some selective reporting of
antibiotic class exposures. Specifically, the smallest study meeting inclusion criteria (36) failed to report effect estimates for 3 of the 7 antibiotic classes (tetracyclines, macrolides, sulfonamides and trimethoprim), and, none of the studies reported on the impact of oxalizidenones, glycopeptides, carbapenems or aminoglycosides.

In summary, and based on the best available evidence, we found that the risk profile for antimicrobial classes as risk factors for community acquired CDI are similar to those described for healthcare-associated disease. In particular, antimicrobial classes with broad spectrum, and potent anti-Gram negative and/or anti-anaerobic activity, including cephalosporins, fluoroquinolones, and clindamycin, are most likely to cause CDI. In contrast, macrolides, penicillins, sulfonamides and trimethoprim, and particularly tetracyclines confer a lower risk of CDI. While community-acquired CDI remains fortunately less common than its healthcare-associated counterpart, we propose that CDI risk represents yet another factor that needs to be factored into the decision to prescribe antimicrobials (and the choice of antimicrobial) in the outpatient setting.
2.7 Acknowledgements

The authors acknowledge Dr. Philippe Vanhems from the Édouard Herriot hospital in Lyon for his guidance and considerate backing of this research partnership.

2.8 Financial Support

This work was supported by Mr. Brown’s Frederick Banting and Charles Best Canada Graduate Scholarship from the Canadian Institutes of Health Research (CIHR). Ms. Khanafer received a grant for her dissertation research from Sanofi Pasteur, France. Dr. Daneman is supported by a Clinician Scientist Salary Award from CIHR. Dr. Fisman received a grant from the Institute of Population and Public Health of the CIHR.

2.9 Potential Conflicts of Interest

All authors: no conflicts.

2.10 References


Chapter 3. The Magnitude and Duration of *Clostridium difficile* Infection Risk After Antibiotic Therapy: a Hospital Cohort Study

### 3.1 Reader’s Note

Although many studies have attempted to quantify antibiotic risk in terms of the degree of cumulative risk over the course of a hospital admission, very few studies have quantified the duration of increased risk associated with antibiotics. In this cohort study, I demonstrate that increased risk associated with antibiotic exposures is concentrated in the period beginning 2 days after antibiotic initiation and ending 5 days after cessation. A version of this paper was accepted in *Public Library of Science ONE* journal and is scheduled for publication on August 26, 2014.

### 3.2 Abstract

Antibiotic therapy is the principal risk factor for *Clostridium difficile* infection (CDI), but little is known about how risk cumulates over the course of therapy and abates after cessation. We prospectively identified CDI cases among adults hospitalized at a tertiary hospital between June 2010 and May 2012. Poisson regression models included covariates for time since admission, age, hospitalization history, disease pressure, and intensive care unit stay; impacts of antibiotic use through time were modeled using 4 measures: receipt on a given day, time since most recent receipt, time since first receipt during a hospitalization, and duration of receipt. Over the 24-month study period, we identified 127 patients with new onset nosocomial CDI (incidence rate per 10,000 patient days \([\text{IR}] = 5.86\)). Of the 4 measures, time since most recent receipt was the strongest independent predictor of CDI incidence. Relative to patients with no prior receipt of antibiotics in the last 30 days \((\text{IR} = 2.95)\), the incidence rate of CDI was 2.41 times higher (95% confidence interval \([\text{CI}] 1.41, 4.13\)) on days when patients received antibiotics and 2.16 times higher...
when patients had receipt in the prior 1-5 days (CI 1.17, 4.00). Risk remained elevated 3-6 days (IR Ratio [IRR] = 3.10, CI 1.73, 5.54) and 7-14 days (IRR = 1.76, CI 0.98, 3.13) following antibiotic initiation. These findings are consistent with studies showing higher risk associated with longer antibiotic use in hospitalized patients, but suggest that the duration of increased risk is shorter than previously thought.

### 3.3 Introduction

*Clostridium difficile* infection (CDI) is a hospital and community-acquired disease that especially impacts elderly hospitalized patients receiving antibiotics (1). CDI incidence in North American hospitals has increased drastically in the last 30 years, and this has been hypothesized to be due to the emergence of new, hyper-virulent strains and to the ubiquity of antibiotic use among hospitalized patients (2). In temperate countries, CDI has a seasonality that follows, with several months delay, that of seasonal prescribing of broad spectrum antibiotics and of pneumonia (3–5).

Antibiotic receipt represents the most important known risk factor for CDI; it is thought to induce CDI risk by denuding the gut of protective bacteria and increasing *C. difficile* spore germination(6,7). Almost all classes of antibiotics have been found to increase CDI risk; certain classes including fluoroquinolones, cephalosporins and clindamycin are thought to have a more potent impact while others, such as tetracyclines, may not change CDI risk at all (8). Several studies have followed patients for periods of 30 to 100 days post-admission and shown that a relatively elevated incidence of CDI exists for patients post-discharge (9–12). Recent studies have used survival analysis incorporating time-varying antibiotic exposures and indicated that increased duration, number, and dosages of antibiotics were associated with increased risk (13,14). Weighted cumulative
exposure models build on survival analysis and stipulate that current risks may be considered as functions of past exposures (15), and may be used to explicitly and flexibly estimate the day-by-day CDI risk during and after antibiotic therapy. As such, our objective was to assess the degree to which risks associated with antimicrobial exposures both cumulate over the course of antimicrobial therapy and abate after cessation.

3.4 Methods

3.4.1 Ethics Statement

Study approval was obtained from the Research Ethics Board of Sunnybrook Health Sciences Centre who waived the need for patient consent because there was no contact with the patients and anonymity was assured.

3.4.2 Study Design and Participants

A case-cohort study design was used to assess the association of antibiotic exposure with the incidence of CDI among patients admitted to Sunnybrook hospital, a large acute care teaching hospital located in Toronto, Canada. The source cohort consisted of all patients over 18 years old, without a previous CDI diagnosis, and hospitalized in an acute care ward at Sunnybrook hospital in the June 1, 2010 to May 31, 2012 period and excluded time spent in the hospital's psychiatry ward.

3.4.3 CDI Case Definition

C. difficile infected patients were prospectively identified by the Infection Prevention and Control (IPC) department via active surveillance during the study period. In accordance with the provincial surveillance definition (16), a CDI case was defined as any hospitalized
patient with either: (a) laboratory confirmation of a positive toxin assay together with diarrhea, or (b) visualization of pseudomembranes on sigmoidoscopy, colonoscopy, or histopathology. For the purposes of surveillance, diarrhea was defined as two or more loose/watery bowel movements in a 24-hour period, which was unusual or different for the patient, and with no other recognized etiology. Case time after the first CDI infection was excluded from the at-risk set, as was all time for patients admitted with non-nosocomial CDI during the study period. Time at risk was restricted to that of hospitalized patients up until the beginning of symptoms of the first disease onset and excluded the first two days of patients’ hospital admissions (patients without previous hospital exposures cannot be considered to have nosocomial acquisition in their first two days of admission according to the provincial CDI definition). Toxin assays at the hospital have been performed by polymerase chain reaction (PCR) since September 2009, which includes the entire study period.

For CDI cases, event time was the number of days from hospital admission to symptom onset, or positive toxin assay for rare cases in which symptom onset was missing. For non-cases, censoring time was the number of days from hospital entry until discharge, study termination, or death.

3.4.4 Antimicrobial Exposure Assessment

Patient antibiotic exposures were drawn from pharmacy dispensing records. We examined for daily receipt of any antibiotic but excluded exposure to metronidazole and oral vancomycin since these may be treatments for CDI. Daily antibiotic receipt was classified according to the Anatomical Therapeutic Chemical Classification System. Only antibiotic classes with exposures during at least 2% of eligible cohort patient-days were analyzed.
individually; preliminary analyses identified penicillins, cephalosporins/ carbapenems, and fluoroquinolones as meeting this criterion. We further subdivided penicillins into broad and narrow spectrum agents and cephalosporins/carbapenems into 1st and 2nd generation cephalosporins, 3rd and 4th generation cephalosporins, and carbapenems, for a total of 8 antibiotic classes that were analyzed individually. We also identified daily receipt of the following 6 infrequently prescribed classes of antibiotics: tetracyclines, nitrofurantoins, sulfanomides and trimethoprim, macrolides and streptogramins, aminoglycosides, and lastly, clindamycin and other lincosamides.

Using the antibiotic classes identified above, we developed two alternative measures of antibiotic exposure: (1) an index representing the number of distinct classes of antibiotics a patient received on a given day (classified as 0, 1 or ≥2), and (2) a categorical antimicrobial risk index which classified patients according to whether they received a high risk antibiotic (defined as receipt of cephalosporins/carbapenems, fluoroquinolones, or clindamycin and other lincosamides), had received a medium risk antibiotic but not a high risk antibiotic (defined as penicillins, sulfanomides and trimethoprim, macrolides and streptogramins, or aminoglycosides), or had received no antibiotics or a low risk antibiotic only (defined as receipt of tetracyclines).

3.4.5  Modeling Time-Varying Antimicrobial Exposures

We created 4 variables based on patients’ unique antibiotic exposure histories: (1) current antibiotic receipt, which was dummy coded as 1 for days when a patient received an antibiotic, and 0 otherwise, (2) the time lapse since the most recent antibiotic use (in days), which ranged from 0 (days of antibiotic receipt) to 31 (lapse > 30 days), (3) time lapse since first antibiotic use, which ranged from 0 (first day of a patients’ first antibiotic
course) to 31 (lapse > 30 days), and (4) the duration of antibiotic use, which ranged from 0 (prior to any antibiotic receipt) to 30 (receipt for each of the previous 30 days). Each of the latter 3 exposures was categorized into approximately equal sized groups.

### 3.4.6 Other Risk Factors

Patient age, sex, hospital pharmacy record, and bed assignment, were obtained from electronic hospital administrative records. Infection pressure was derived by calculating the number of diagnosed infectious cases of CDI in each ward of the hospital at noon of each day using patient location records. Both nosocomial and non-nosocomial cases could contribute to daily measures of infection pressure. The infectious period of a given case was defined as starting on the day after symptom onset to 14 days after the positive test associated with case detection (17). We also measured the use of antacids, laxatives, feeding tube, and whether a patient had stayed in an intensive care unit (ICU).

### 3.4.7 Statistical Analysis

For bivariate analyses, we compared characteristics of the at-risk period of cases with a 10% simple random sample of control patients. Two-sided p-values were assessed with Pearson’s chi-squared test for categorical variables and with the Wilcoxon rank-sum test for continuous variables. For reporting incidence rates and for the regression models, we weighted all of the control patients’ follow-up times by 10 and all of the case patients’ follow up times by 1, so as to reflect rates from the original, complete, cohort (18).

To estimate the impact of antibiotic exposures on CDI risk, we developed weighted Poisson regression models that aimed to predict the time elapsed from hospital admission to the occurrence of a first CDI infection. Our data was structured in counting process
format with one record for each patient-day. For each of the 4 antibiotic exposure covariates, two models were fitted to the data. The 4 unadjusted models included no covariates other than antibiotic exposures; incidence rate ratios were estimated using Poisson regression. The 4 adjusted models included 6 potential confounders: time since admission (modeled as a b-spline with knots at 5 and 15 days), patient age (classified as <45, 45 to <65, ≥65 years), sex, number of previous hospital admissions (classified as 0, 1, ≥2), patient-days of infection pressure in the past 10 days, and ICU admission in the past 10 days. The number of adjustment factors was restricted in order to ensure at least 10 events per covariate (19), and the selection of covariates was based on established associations with CDI risk (13,20). For the adjusted models, intra-patient correlation was accounted for using the generalized estimating equation approach (21). Statistical fit for all models was assessed using Akaike’s Information Criterion (22).

Subsequently, we used the best fitting of the 4 antibiotic exposure covariates to determine risk associated with the two risk indexes and for each of the 8 antibiotic classes. For unadjusted and adjusted estimates of the antibiotic-specific risks, antibiotic exposure adjustment variables were derived which measured exposure to any other antibiotic without exposure to the antibiotic in question.

Analyses were conducted using R statistical software (v3.0.2); the glm and geeglm (23) functions were used for the unadjusted and adjusted statistical models, respectively. R statistical software analysis code is provided in Appendix S1.
3.4.8 Sensitivity Analyses

In order to assess the potential impact of uncertainty of diagnostic timing on the analyses, we conducted a sensitivity analysis using positive *C. difficile* PCR test date rather than symptom onset date to define the outcome timing. Also, because pre-admission antibiotic exposure information was not available, we conducted an additional sensitivity analysis excluding cases and patient time in the first 10 days of each hospitalization to investigate the impact of exposure history incompleteness.

3.5 Results

3.5.1 Description of Cohort

Over the two-year study period, and before exclusion of ineligible patients, a total of 47,241 patients were identified as having been admitted to Sunnybrook Health Science Centre (Figure 3.1). Of these, 412 were diagnosed with CDI; after exclusion of patients with non-nosocomial CDI, and patients with onset of CDI while out of hospital, or within the first two days of an admission, 127 nosocomial case patients remained. After removal of ineligible non-case patients, the 10% control cohort selection consisted of 1,940 patients. The incidence of CDI in the cohort was 5.86 per 10,000 days of follow-up (127/216,978).
3.5.2 Demographic and Exposure Characteristics of Cases and Controls

The median age of cases (72.0) was almost 5 years older than that of controls (67.3, p=0.14). Cases had a higher rate of exposure to other symptomatic CDI patients in their own ward (median, 20 per 100 person-days versus 0 per 100 person-days, p=0.12). About half (50.4%) of case patients spent time in the ICU during their risk period compared to 19.5% of controls (p<0.001).

A larger proportion of case patients received antibiotics during their risk period relative to controls (86.6% vs 49.5%, p < 0.001). The majority (71.7%) of cases received a cephalosporin class of antibiotic compared to 33.1% of controls (p<0.001). Similarly, penicillins and fluoroquinolones were more likely to be prescribed among case patients compared to controls. Among case patients, median symptom onset date was 9 days after admission (IQR: 5, 17 days, Figure 3.2).
Figure 3.2. Symptom Onset Time of Incident *Clostridium difficile* Infection (N=127), Sunnybrook Hospital, Toronto, Canada, June 2010 to May 2012.
<table>
<thead>
<tr>
<th></th>
<th>Incident Cases (N=127)</th>
<th>Controls (N=1940)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (IQR), y</td>
<td>72.0 (57.6-79.8)</td>
<td>67.3 (53.2-79.2)</td>
<td>0.14</td>
</tr>
<tr>
<td>Male gender</td>
<td>69 (54.3)</td>
<td>989 (51.0)</td>
<td>0.52</td>
</tr>
<tr>
<td>Admissions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>76 (59.8)</td>
<td>1494 (77.0)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>2</td>
<td>30 (23.6)</td>
<td>302 (15.6)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>≥3</td>
<td>21 (16.5)</td>
<td>144 (7.4)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ICU stay</td>
<td>64 (50.4)</td>
<td>379 (19.5)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Disease pressure&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>median per 100 patient-days (IQR)</td>
<td>20.0 (0.0-50.0)</td>
<td>0.0 (0.0-50.0)</td>
<td>0.12</td>
</tr>
<tr>
<td>Days of antibiotic exposure,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>median per 100 patient-days (IQR)</td>
<td>46.2 (25.0-76.0)</td>
<td>0.0 (0.0-60.0)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Antibiotic exposure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>110 (86.6)</td>
<td>961 (49.5)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Penicillins</td>
<td>38 (29.9)</td>
<td>220 (11.3)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Broad-spectrum</td>
<td>29 (22.8)</td>
<td>191 (9.8)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Narrow-spectrum</td>
<td>12 (9.4)</td>
<td>43 (2.2)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Cephalosporins/carbapenems</td>
<td>91 (71.7)</td>
<td>642 (33.1)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; &amp; 2&lt;sup&gt;nd&lt;/sup&gt; generation</td>
<td>44 (34.6)</td>
<td>388 (20.0)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; &amp; 4&lt;sup&gt;th&lt;/sup&gt; generation</td>
<td>59 (46.5)</td>
<td>326 (16.8)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Carbapenems</td>
<td>11 (8.7)</td>
<td>41 (2.1)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Fluoroquinolones</td>
<td>49 (38.6)</td>
<td>381 (19.6)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>IV vancomycin</td>
<td>28 (22.0)</td>
<td>107 (5.5)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Other exposures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antacids (H2 inhibitors)</td>
<td>96 (75.6)</td>
<td>1268 (65.4)</td>
<td>0.024</td>
</tr>
<tr>
<td>Laxatives</td>
<td>91 (71.7)</td>
<td>1187 (61.2)</td>
<td>0.024</td>
</tr>
<tr>
<td>Feeding tube</td>
<td>53 (41.7)</td>
<td>263 (13.6)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Abbreviations: ICU, intensive care unit; IQR, interquartile range

<sup>a</sup> equal to the number of patients diagnosed with CDI in the same ward as a given patient each day.

<sup>b</sup> 2 degree of freedom Pearson’s Chi-square test.
### Table 3.2. Timing and Magnitude of CDI Risk Associated with 4 Different Antibiotic Exposure Measures

<table>
<thead>
<tr>
<th>CDI Cases (N)</th>
<th>Follow-up (days)</th>
<th>Unadjusted</th>
<th>Adjusted&lt;sup&gt;c&lt;/sup&gt;</th>
<th>( \Delta \text{AIC}) &lt;sup&gt;b&lt;/sup&gt;</th>
<th>( \Delta \text{AIC}) &lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>IRR&lt;sup&gt;a&lt;/sup&gt;</td>
<td>IRR (95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antibiotic use on current day</td>
<td></td>
<td></td>
<td></td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>No</td>
<td>63</td>
<td>141019</td>
<td>4.47</td>
<td>Reference</td>
<td>1.89 (1.33, 2.67)</td>
</tr>
<tr>
<td>Yes</td>
<td>64</td>
<td>75959</td>
<td>8.43</td>
<td>2.86 (1.73, 4.72)</td>
<td>1.79 (1.24, 2.59)</td>
</tr>
<tr>
<td>Time since end antibiotic therapy (d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 (current receipt)</td>
<td>64</td>
<td>75959</td>
<td>8.43</td>
<td>2.86 (1.73, 4.72)</td>
<td>2.41 (1.41, 4.13)</td>
</tr>
<tr>
<td>1-5</td>
<td>30</td>
<td>36950</td>
<td>8.12</td>
<td>2.75 (1.56, 4.85)</td>
<td>2.16 (1.17, 4.00)</td>
</tr>
<tr>
<td>6-30</td>
<td>13</td>
<td>36288</td>
<td>3.58</td>
<td>1.21 (0.60, 2.44)</td>
<td>0.98 (0.48, 2.00)</td>
</tr>
<tr>
<td>&gt; 30, or no antibiotic use</td>
<td>20</td>
<td>67781</td>
<td>2.95</td>
<td>Reference</td>
<td>-9.5</td>
</tr>
<tr>
<td>Time since start of first antibiotic (d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-2</td>
<td>7</td>
<td>21060</td>
<td>3.32</td>
<td>1.08 (0.47, 2.49)</td>
<td>1.39 (0.56, 3.46)</td>
</tr>
<tr>
<td>3-6</td>
<td>40</td>
<td>38007</td>
<td>10.52</td>
<td>3.43 (2.11, 5.59)</td>
<td>3.10 (1.73, 5.54)</td>
</tr>
<tr>
<td>7-14</td>
<td>33</td>
<td>37530</td>
<td>8.79</td>
<td>2.87 (1.72, 4.77)</td>
<td>1.76 (0.98, 3.13)</td>
</tr>
<tr>
<td>15-30</td>
<td>20</td>
<td>32313</td>
<td>6.19</td>
<td>2.02 (1.13, 3.60)</td>
<td>1.56 (0.84, 2.90)</td>
</tr>
<tr>
<td>&gt; 30, or no antibiotic use</td>
<td>27</td>
<td>88068</td>
<td>3.07</td>
<td>Reference</td>
<td>-14.6</td>
</tr>
<tr>
<td>Cumulative duration of antibiotic use (d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 (no prior receipt)</td>
<td>20</td>
<td>67781</td>
<td>2.95</td>
<td>Reference</td>
<td>1.87 (1.04, 3.37)</td>
</tr>
<tr>
<td>1-3</td>
<td>25</td>
<td>45201</td>
<td>5.53</td>
<td>2.74 (1.55, 4.87)</td>
<td>2.27 (1.24, 4.16)</td>
</tr>
<tr>
<td>4-6</td>
<td>28</td>
<td>34596</td>
<td>8.09</td>
<td>2.64 (1.49, 4.68)</td>
<td>2.10 (1.12, 3.94)</td>
</tr>
<tr>
<td>7-11</td>
<td>28</td>
<td>35968</td>
<td>7.78</td>
<td>2.64 (1.47, 4.72)</td>
<td>2.84 (1.39, 5.81)</td>
</tr>
<tr>
<td>&gt;11</td>
<td>26</td>
<td>33432</td>
<td>7.78</td>
<td>2.64 (1.47, 4.72)</td>
<td>2.84 (1.39, 5.81)</td>
</tr>
</tbody>
</table>

Abbreviations: AIC, Akaike’s Information Criterion; CDI, *Clostridium difficile* infection; CI, confidence interval; d, days; IR, incidence rate; IRR, incidence rate ratio.

<sup>a</sup> Incidence rate, per 10,000 patient-days.

<sup>b</sup> The difference in AIC relative to the reference model (current antibiotic use): negative numbers denote an improvement in fit. \( \Delta \text{AIC} < -2 \) was considered a statistically significant improvement in fit at \( p < 0.05 \).

<sup>c</sup> Each of the 4 adjusted models had 7 covariates: one antibiotic exposure variable (as described in the table), in addition to the following 6 potential confounders: time since hospital admission, age, gender, number of previous hospital admissions, infection pressure, and current or prior ICU admission.
Table 3.3. CDI Risk Associated with Antimicrobial Exposures During and Within 5 days of the End of Antimicrobial Therapy, for Antibiotic Risk Indexes and Specific Antibiotic Exposures.

<table>
<thead>
<tr>
<th>Exposure in the preceding 5d</th>
<th>CDI Cases (N)</th>
<th>Follow-up (days)</th>
<th>Unadjusted</th>
<th>Adjusted&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>IR&lt;sup&gt;a&lt;/sup&gt;</td>
<td>IR Ratio (95% CI)</td>
</tr>
<tr>
<td>Any antibiotic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>33</td>
<td>104069</td>
<td>3.2</td>
<td>Reference</td>
</tr>
<tr>
<td>Yes</td>
<td>94</td>
<td>112909</td>
<td>8.3</td>
<td>2.63 (1.77, 3.90)</td>
</tr>
<tr>
<td>Number of antibiotics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>33</td>
<td>104069</td>
<td>3.2</td>
<td>Reference</td>
</tr>
<tr>
<td>1</td>
<td>64</td>
<td>73593</td>
<td>8.7</td>
<td>2.74 (1.80, 4.17)</td>
</tr>
<tr>
<td>≥ 2</td>
<td>30</td>
<td>39316</td>
<td>7.6</td>
<td>2.41 (1.47, 3.95)</td>
</tr>
<tr>
<td>Antibiotic risk index</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None or low-risk</td>
<td>35</td>
<td>106396</td>
<td>3.3</td>
<td>Reference</td>
</tr>
<tr>
<td>Medium-risk</td>
<td>12</td>
<td>19256</td>
<td>6.2</td>
<td>1.89 (1.33, 2.67)</td>
</tr>
<tr>
<td>High-risk</td>
<td>80</td>
<td>91326</td>
<td>8.8</td>
<td>2.66 (1.79, 3.96)</td>
</tr>
<tr>
<td>Class of antibiotic&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Penicillins</td>
<td>24</td>
<td>25103</td>
<td>9.5</td>
<td>3.02 (1.78, 5.10)</td>
</tr>
<tr>
<td>Broad-spectrum</td>
<td>16</td>
<td>18876</td>
<td>8.5</td>
<td>2.67 (1.47, 4.86)</td>
</tr>
<tr>
<td>Narrow-spectrum</td>
<td>8</td>
<td>6857</td>
<td>11.7</td>
<td>3.68 (1.70, 7.97)</td>
</tr>
<tr>
<td>Cephalosporins</td>
<td>61</td>
<td>66330</td>
<td>9.2</td>
<td>2.90 (1.90, 4.43)</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; &amp; 2&lt;sup&gt;nd&lt;/sup&gt; generation</td>
<td>31</td>
<td>38508</td>
<td>8.0</td>
<td>2.54 (1.55, 4.15)</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; &amp; 4&lt;sup&gt;th&lt;/sup&gt; generation</td>
<td>32</td>
<td>27092</td>
<td>11.8</td>
<td>3.72 (2.29, 6.06)</td>
</tr>
<tr>
<td>Carbapenems</td>
<td>6</td>
<td>5962</td>
<td>10.1</td>
<td>3.17 (1.33, 7.57)</td>
</tr>
<tr>
<td>Fluoroquinolones</td>
<td>24</td>
<td>30286</td>
<td>7.9</td>
<td>2.50 (1.48, 4.23)</td>
</tr>
<tr>
<td>IV vancomycin</td>
<td>16</td>
<td>11883</td>
<td>13.5</td>
<td>4.25 (2.34, 7.71)</td>
</tr>
</tbody>
</table>

Abbreviations: AIC, Akaike’s Information Criterion; CDI, *Clostridium difficile* infection; CI, confidence interval; d, days; IR, incidence rate; IRR, incidence rate ratio.

<sup>a</sup> Incidence rate, per 10,000 patient-days.

<sup>b</sup>The difference in AIC relative to the reference model (receipt of any antibiotic in the previous 5 days): negative numbers denote an improvement in fit. Δ AIC < −2 was considered a statistically significant improvement in fit at p < 0.05.

<sup>c</sup>Adjusted for time since hospital admission, age, gender, number of previous hospital admissions, infection pressure, and current or prior ICU admission.

<sup>d</sup>Each antibiotic group was assessed in a separate model; the reference group for each relative risk estimate was no receipt of antibiotics in the last 5 days.
3.5.3 Risk Associated with Antibiotic Exposures

The incidence of CDI when patients received antibiotics was 8.43 per 10,000 days (64/75,959) compared to 4.47 (63/141,019) when patients did not receive antibiotics (Table 3.2); it follows that the incidence rate ratio (IRR) associated with current antibiotic exposure was 1.89 (95% confidence interval [CI] 1.33, 2.67). Adjustment for time since admission, patient age, gender, number of previous hospital admissions, disease pressure, and current or prior ICU admission reduced the IRR slightly, to 1.79 (95% CI 1.24, 2.59).

Figure 3.3. The Magnitude and Duration of Clostridium difficile Infection Risk After Antibiotic Therapy, Sunnybrook Hospital, Toronto, Canada, June 2010 to May 2012. Among inpatients, the incidence of Clostridium difficile infection was highest in the period 3 to 14 days after the start of antibiotic therapy, during antibiotic therapy, and within 5 days of the end of antibiotic therapy. * Includes patients without any identified antibiotic use.

Among the 4 antibiotic exposure measures, the time since last antibiotic use was the most important independent predictor of CDI onset, and yielded a statistically significant improvement in the prediction of CDI onset (Δ adjusted AIC = -4.6). The incidence of CDI (per 10,000 days) was 2.95 when patients had received no antibiotics in
the previous 30 days (reference), 8.43 when patients were currently receiving antibiotics (adjusted IRR: 2.41, 95% CI 1.41, 4.13), 8.12 when patients had recent receipt in the last 1-5 days (adjusted IRR: 2.16, 95% CI 1.19, 3.32), and 3.58 when patients had received antibiotics in the last 6-30 days (adjusted IRR: 0.98, 95% CI 0.48, 2.00, Figure 3.3).

The time elapsed since the start of the first antibiotic exposure in a given hospitalization also yielded an improved prediction of the timing of CDI onset ($\Delta$ adjusted AIC = -0.7). The observed association was highest 3-6 days (adjusted IRR = 3.10, 95% CI 1.73-5.54) and 7-14 days (adjusted IRR = 1.76, 95% CI 0.98-3.13) after the start of antibiotics. Although the duration of antibiotic exposure was associated with increased risk, the duration of antibiotic exposure variable yielded a worse statistical fit relative to the adjusted model with current antibiotic use ($\Delta$ adjusted AIC = 5.1).

### 3.5.4 Sensitivity Analyses

For all remaining analyses including the sensitivity analyses and the investigation of antibiotic class-specific effects, we categorized antibiotic exposure history as receipt of antibiotics in the last 5 days (which included current receipt), or no receipt in the last 5 days. The adjusted risk was 2.35 (95% CI 1.53, 3.60) for patients that had received an antibiotic in the previous 5 days (Table 3.3). We conducted 2 different sensitivity analyses; neither impacted the estimated risk substantially. In the first sensitivity analysis, we considered the impact of restricting the dataset to follow-up $\geq$ 10 days after admission, so that complete information on antibiotic history was known for a larger proportion of patients. The adjusted IRR for risk extended 5 days beyond the end of antibiotic use was similar to that of the full cohort (2.23, 95% CI 1.21, 4.13). We also considered a sensitivity analysis in which we varied the assignment of CDI event date; when positive test date
rather than symptom onset was used to define the outcome, the number of eligible cases increased from 127 to 130. The adjusted incidence rate for antibiotic use in the last 5 days was similar, at 2.79 (95% CI 1.47, 5.27).

3.5.5 Antibiotic Risk According to Antimicrobial Classes

In order to consider differences in the level of risk among antibiotic users, we considered risk among patients with exposure to various combinations of antibiotics and to specific antibiotics (Table 3.3). The adjusted risk associated with CDI was similar when patients either received a single class of antibiotic (IRR = 2.49, 95% CI 1.59, 3.92) or when patients received multiple classes of antibiotics (IRR = 2.09, 95% CI 1.23, 3.55). For our antibiotic risk index variable, which was based on established associations of antibiotics with CDI risk, the adjusted risk was 1.79, (95% CI 1.24, 2.59) for exclusive low-risk antibiotics and was 2.43 (95% CI 1.59, 3.74) for high-risk antibiotics.

For each of the 8 antibiotic classes and subclasses for which there was sufficient exposure for individual analysis (>2% of subcohort patient-time), we measured the risk of the given antibiotic taken alone or in combination with other antibiotics, relative to no antibiotic exposure. In adjusted analyses, all 3 of the most prescribed antimicrobial classes demonstrated similarly large hazard ratios. Of the subclasses, 3rd & 4th generation cephalosporins taken alone or in combination, had higher risk in comparison to other antibiotic classes (adjusted IRR=3.40, 95% CI 2.02, 5.72).

3.6 Discussion

Our observational study of 127 CDI cases and 2 years of follow-up on patients at a large tertiary hospital found that: (1) in-hospital CDI risk was highest 3 to 14 days after the start
of the first antibiotic course, (2) elevated CDI risk persisted after the end of antimicrobial therapy, and (3) patients with longer antibiotic courses were at higher risk of developing CDI than patients with shorter courses, but even short courses and single doses of antibiotics incur a substantial risk of inducing CDI.

In a case-control study of 337 patients with hospital-associated CDI, risk was found to be relatively constant both during antibiotic use and for a period of 30 days after cessation (24), diminishing more than 30 days after the end of antimicrobial exposure. Our study population was restricted to inpatients that may have developed CDI more rapidly than outpatients and patients discharged from hospital.

Our results showed that measuring the time since last antibiotic use is the most predictive metric for quantifying risk for a patient, whereby risk during and within 5 days of the end of antibiotic therapy was the most elevated. Since colonization resistance is thought to be greatly diminished both during and for a period after the end of antibiotic use, our findings support the importance of this mechanism and reflect findings from animal and in vitro models of colonization resistance as it relates to CDI (25,26). Inferring from the low risk for 2 days after the start of antimicrobial use, our models provide empirical evidence that the incubation period of CDI is at least 48h (27).

A limitation of our study was our lack of information on outpatient antimicrobial exposures prior to patient hospitalization, since our exposure-free reference group used for calculating hazard ratios could have included patients exposed to antibiotics prior to admission. However, our findings were robust in sensitivity analyses considering subsets of patients with more prolonged hospitalization and therefore more complete antibiotic exposure histories. Further, our study lacked information on post-discharge *C. difficile*
diagnoses. Considering discharged patients as censored surmounted this limitation, but means that our results are most generalizable to acute, hospital-onset CDI. Furthermore, clinical teams are aware of antimicrobial exposures which may prompt diagnostic testing for CDI [27], and this surveillance bias may lead to an overestimation of the risk associated with antibiotic exposure. Finally, we had no data on *C. difficile* colonization status and timing of acquisition of the organism, which would be expected to influence the time lapse between antibiotic exposure and disease onset.

Antibiotic use is the most important risk factor for CDI, and a substantial amount of research has considered the risk of CDI associated with different antibiotic exposures. In this study of the association between antibiotic exposures and CDI risk in hospitalized adults, we focused on the timing of increased risk; we found that risk appears 3 days after the onset of antibiotic use, and continues for a period of 5 days after the end of antibiotics, and is relatively unimportant thereafter. Further research may consider how different antibiotics may induce different time-varying risks in order to differentiate antibiotic impacts and improve patient outcomes.

### 3.7 Author Contributions

Conceived and designed the experiments: KAB ND DNF. Performed the experiments: KAB. Analyzed the data: KAB RM. Wrote the paper: KAB DNF. The corresponding author had full access to all the data and had final responsibility for the decision to submit for publication.
3.8 References


3.9 Appendix 1. Unadjusted and Adjusted Estimates of Other Risk Factors

In addition to antibiotic exposure, patient age, sex, hospitalization number, ICU stay, and *Clostridium difficile* disease pressure and the number of days since admission were included in the adjustment model. Patient age was significantly associated with risk. The incidence of CDI among patients 18-44 years was 3.91 (11/28199) per 10,000 patient-days compared to 5.62 (34/60445, adjusted IRR 1.47, 95% CI 0.67, 3.25) among patients 45-64 years and 6.39 (82/128414, adjusted IRR 1.82, 95% CI 0.87, 3.78) among patients ≥65 years. Hospitalization number was also an important risk factor. Patients presenting for their second hospitalization during the study period were 1.82 times (adjusted IRR, 95% CI 1.14, 2.91) more likely to develop disease and patients presenting 3 or more times were 2.47 (adjusted IRR, 95% CI 1.44, 4.25) more likely to develop disease during these latter visits. 42% of cases developed CDI during or following an ICU admission, and ICU admission in the previous 10 days was significantly associated with patient risk (adjusted IRR 1.63, 95% CI 1.08, 2.45). There was no association between *Clostridium difficile* disease pressure and CDI risk. The incidence of CDI was 6.09 per 10,000 patient-days among patients without any disease pressure, 4.84 per 10,000 patient-days among patients with 1-5 days of disease pressure and 7.02 per 10,000 patient days (adjusted IRR 0.69, 95% CI 0.45, 1.07) among patients with 6 or more days of disease pressure (adjusted IRR 1.00, 95% CI 0.60, 1.66).
Table 3.4. CDI Risk Associated with Exposures Retained in the Adjusted Analyses*.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>CDI Cases (N)</th>
<th>Follow-up (days)</th>
<th>Unadjusted IR Ratio (95% CI)</th>
<th>Adjusted IR Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any antibiotic in the last 5 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>33</td>
<td>104069</td>
<td>3.17 Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Yes</td>
<td>94</td>
<td>112909</td>
<td>8.33 2.63 (1.77, 3.9)</td>
<td>2.37 (1.54, 3.63)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;45</td>
<td>11</td>
<td>28119</td>
<td>3.91 Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>45-64</td>
<td>34</td>
<td>60445</td>
<td>5.62 1.44 (0.73, 2.84)</td>
<td>1.47 (0.67, 3.25)</td>
</tr>
<tr>
<td>≥65 years</td>
<td>82</td>
<td>128414</td>
<td>6.39 1.63 (0.87, 3.06)</td>
<td>1.82 (0.87, 3.78)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>58</td>
<td>98739</td>
<td>5.87 Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Male</td>
<td>69</td>
<td>118239</td>
<td>5.84 0.99 (0.70, 1.41)</td>
<td>0.93 (0.62, 1.38)</td>
</tr>
<tr>
<td>Hospitalization Number</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>76</td>
<td>160513</td>
<td>4.73 Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>37422</td>
<td>8.02 1.69 (1.11, 2.58)</td>
<td>1.82 (1.14, 2.91)</td>
</tr>
<tr>
<td>≥3</td>
<td>21</td>
<td>19043</td>
<td>11.03 2.33 (1.44, 3.78)</td>
<td>2.47 (1.44, 4.25)</td>
</tr>
<tr>
<td>ICU stay in the last 10 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>74</td>
<td>154163</td>
<td>4.80 Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Yes</td>
<td>53</td>
<td>62815</td>
<td>8.44 1.76 (1.24, 2.5)</td>
<td>1.63 (1.08, 2.45)</td>
</tr>
<tr>
<td>* Clostridium difficile Disease Pressure in the last 10 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>69</td>
<td>113262</td>
<td>6.09 Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>1-5 patient-days</td>
<td>33</td>
<td>68126</td>
<td>4.84 0.80 (0.53, 1.20)</td>
<td>0.69 (0.45, 1.07)</td>
</tr>
<tr>
<td>≥6 patient-days</td>
<td>25</td>
<td>35590</td>
<td>7.02 1.15 (0.73, 1.82)</td>
<td>1.00 (0.60, 1.66)</td>
</tr>
</tbody>
</table>

Abbreviations: CDI, Clostridium difficile infection; CI, confidence interval; d, days; IR, incidence rate; IRR, incidence rate ratio.

* Note that in addition to the risk factors above, time since admission was also included as a beta-spline with knots at 5 and 15 days. Also, Clostridium difficile Disease Pressure in the last 10 days was included as a linear predictor in the final analyses, while for the table above it is presented as categorized into bins of 0, 1 to 5, and ≥6 patient-days of exposure.
Appendix 2. R Analysis Code For Tables and Poisson Regression Models

# March 3, 2014

# R code for producing the principal analyses of The magnitude and duration of Clostridium difficile infection risk after antibiotic therapy: a hospital cohort study.

# For this code to run, one requires the daily.csv dataset
# (a dataset of records for each patient-day). The authors are
# currently considering depositing the dataset on Dryad.

```{r}
# Load custom functions

getIRs<-function(y,x=NULL,weights,r=2){
  if(is.null(x)){ x<-rep(1,length(y)) }
  n <- tapply(y*weights,x,sum)
  d <- tapply(weights,x,sum)
  cbind(N=n,PT=d,IR=round(n/d*10000,2))
}

getIRs.f<-function(f,data,r=2){
  tbl<-xtabs(f,data=data)
  n<-tbl[,2]
  d<-tbl[,1]
  cbind(N=n,PT=d,IR=round(n/d*10000,2))
}

# Create a 2 column table of bivariate statistics;
# Specify:
# data = a data frame (each column will be a row in the table)
# group = the grouping vector to specify columns for the comparisons
# statistic = a vector of "p" or "q"
#   p = proportion / chisquare test
#   q = median (p25, p75) / wilcoxon rank-sum / kruskal wallis
# depends on percentn
# depends on iqr

table1<-function(data,group,statistic){
  varnames<-colnames(data)
  n<-length(statistic)
  group<-as.factor(group)
  l<-levels(group)
  g<-length(l)
  x<-matrix(data=NA,nrow=n,ncol=g+1)
  rownames(x)<-paste(varnames,statistic)
  colnames(x)<-c(l,"p.value")

  for(i in 1:n){
    for(j in 1:g){
      v<-data[group==l[j],i]
      if(statistic[i]=="q")
        x[i,j]<-iqr(v)
      if(statistic[i]=="p")
        x[i,j]<-percentn(v)
    }
  }
  if(statistic[i] %in% c("q"))
    p<-kruskal.test(x=data[,i],g=group)$p.value
```

```
```
if(statistic[i] %in% c("p"))
  p <- chisq.test(table(data[,i]>0,group))$p.value

if(p<0.0005)
  p.char<-"<0.001"
else if(p>0.99)
  p.char<-"0.99"
else
  p.char<-round(p,3)

x[i,g+1]<-p.char

x

iqr<-function(x,r=1){
  f<-paste("%.",r,"f",sep="")
  paste(sprintf(f,median(x))," (",
  sprintf(f,quantile(x,.25)),"-",
  sprintf(f,quantile(x,.75))," )",sep="")
}

percentn<-function(x,r=1){
  f<-paste("%.",r,"f",sep="")
  paste( sum(x>0) ," (", sprintf (f, mean(x>0) * 100 ), ")",sep="")
}

# create n x m tables with n and %
# m is the margin to use for the % (1=row / 2=col)
percentn.t<-function(x,y,r=1,m=2){
  f<-paste("%.",r,"f",sep="")
  x<-as.character(x)
  y<-as.character(y)
  tab. <-table(x,y)
  l <- length(x)
  r <- dim(tab.)[1]
  c <- dim(tab.)[2]
  margins <- apply(tab.,m,sum)
  pct <- paste( tab. , " (", sprintf (f, pct * 100 ), ")",sep="")
  out <- matrix(v,nrow=r)
  rownames(out) <- rownames(tab.);
  colnames(out) <- colnames(tab.);
  out
}

percentn.f<-function(x,r=1){
  f<-paste("%.",r,"f",sep="")
  x<-as.character(x); x<-as.factor(x);
  levs<-sort(levels(x));
  l<-length(levs);
  out<-as.character(1:l)
  for(i in 1:l)
    out[i]<-paste( sum(x==levs[i]) ," (", sprintf (f, mean(x==levs[i]) * 100 ), ")",sep="")
  out
}

citable<-function(x, dash=" to "){
  old.col<-ncol(x)
  new.col<-old.col/3
  X<-matrix("",ncol=new.col,nrow=nrow(x))
  rownames(X)<-rownames(x)
for(i in 1:new.col){
    j<-(i-1)*3 + 1;
    X[,i]<-paste(x[,j], " (", x[,j+1], dash, x[,j+2], ")", sep=""
    X[,i]<-ifelse(X[,i]==paste('NA',' (', 'NA', dash, 'NA', " ), sep="'",',X[,i])
}

# Retrieve the effect estimates and the 95% confidence intervals from various regression models
# If e = T then return exp(ef)
# round = the number of decimal places to use for returning the answer

eff.95<-function(model, e=F, round=2){
    eff<-coef(model)
    se<-sqrt(diag(vcov(model)))
    lower<-eff-1.96*se
    upper<-eff+1.96*se
    out<-cbind(eff,lower,upper)
    if(e==T) out<-exp(out)
    round(out,round)
}

eff.95.gee<-function(model,e=F,round=2){
    eff<-summary(model)[7]$coefficients[,1]
    se<-summary(model)[7]$coefficients[,4]
    lower<-eff-1.96*se
    upper<-eff+1.96*se
    out<-cbind(eff,lower,upper)
    if(e==T) out<-exp(out)
    round(out,round)
}

eff.95.geeglm<-function(model,e=F,round=2){
    eff<-coef(model)
    se<-sqrt(diag(model$geese$vbeta))
    lower<-eff-1.96*se
    upper<-eff+1.96*se
    out<-cbind(eff,lower,upper)
    if(e==T) out<-exp(out)
    round(out,round)
}

going<-function(m){
    cbind(eff.95(m,T), AIC=round(AIC(m),2));
}

generate<-function(m){
    cbind(eff.95.geeglm(m,T), QIC=round(QIC2(m),2));
}

# QIC for GEE models
# Daniel J. Hocking
# QIC2 = function(model.R) {
library(MASS)
model.indep = update(model.R, corstr = "independence")
# Quasilikelihood
mu.R = model.R$fitted.values
y = model.R$y
type = family(model.R)$family
quasi.R = switch(type,
    poisson = sum((y*log(mu.R)) - mu.R),
    gaussian = sum(((y - mu.R)^2)/-2),
    binomial = sum(y*log(mu.R/(1 - mu.R)) + log(1 - mu.R)),
}
\[
\Gamma = \text{sum}\left(-\frac{y}{(mu.R - \log(mu.R))}\right)
\]

# Trace Term (penalty for model complexity)
omegai = ginv(model.indep$geese$vbeta.naiv) # Omega-hat(I) via Moore-Penrose
# AIinverse = solve(model.indep$geese$vbeta.naiv) # solve via indentity
Vr = model.R$geese$vbeta
px = length(mu.R) # number non-redundant columns in design matrix
# QIC
QIC = 2*(trace.R - quasi.R) # [EDIT: original post was missing '*']
#QICu = (-2)*quasi.R + 2*px # Approximation assuming model structured correctly
output = c(QIC, quasi.R, trace.R, px)
names(output) = c('QIC', 'Quasi Lik', 'Trace', 'px')
return(output)

```{r Figure 2 - Symptom Onset (days from admission)}
par(mfrow=c(1,1),mar=c(4,4,1,2))
hist(pmin(daily$t0_hospit,50)[daily$cdi_event==1],c(-1:50,999)+.5,freq=T,xlim=c(0,52),
main='',ylab='Symptom Onset (days from admission)',col='black',border='white', axes=F)
axis(1,c(2,1:5*10),c(2,1:4*10,'>50'))
axis(2,0:6*2)
```
```{r Table1}
# Table 1. Comparison of at-risk patients, and person time for case patients versus
# control patients. Note that *type* variable depicts patients by whether they acquired
# CDI or not (type = 1 CDI, type = 2 noCDI). Note that *cdi_event* variable is 1 only
# on the day that patients acquire CDI. 127 have event = 1
# Cases
describe(daily[daily$cases==1,"pid"]) # total follow-up time among cases
describe(daily[daily$sc==1 & daily$cases==1,"pid"]) # follow-up among subcohort-selected cases
# Controls
describe(daily[daily$cases==0,"pid"]) # N=1940, 21,433 days
# Variables for Table 1
v.id<-c("pid","t0","t0_hospit","cases")
v.cont<-c("age","dp2.100","medany.100")
v.prop<-c("male","dp2","icu",paste("med",c("any",3,19,20,4,22,21,23,8,24,11:12,18),sep=""))

```
# Combine and create final table
y<-rbind(y.mean, y.max)
colnames(y)[1:2]<-c("Controls (n=1940)","Cases (n=127)"); y<-y[,c(2:1,3)]; y

# Put the table in the correct order
o<-c(1,4,20:22,6,2,3,7:19)
y<-y[o,

# Output table
y
print(xtable(y),type="html",file="Table1.html")
```
```{r Create Table 2
- incidence rates for diff ABx exposures}

# Overall IR
t0<-with(daily,getIRs(cdi_event,weights=weight)); t0

# Incidence rates in the exposure categories;
t1<-with(daily,getIRs(cdi_event,medany,weight))
t2<-with(daily,getIRs(cdi_event,medany_stop_cat,weight))
t3<-with(daily,getIRs(cdi_event,medany_start_cat,weight))
t4<-with(daily,getIRs(cdi_event,medany_sum_cat,weight))

irs<-rbind(t1,t2,t3,t4); irs

# Unadjusted models;
m.un<-glm(cdi_event ~ 1,  family=poisson("log"), data=daily, weights=weight);
m <- update(m.un, . ~ medany + . )
u1<-getUn(m)
m <- update(m.un, . ~ medany_stop_cat + . )
u2<-getUn(m)
m <- update(m.un, . ~ medany_start_cat + . )
u3<-getUn(m)
u4<-getUn(m)

un<-rbind(u1,u2,u3,u4); un

# Adjusted models
m.adj<-geeglm(cdi_event ~ bx.t0_hospit + as.factor(age4) + male + as.factor(n_hospit2) + I(icu_stop<=10) + pmin(dp2__10,10), id=pid, family="poisson", corstr="exchangeable",data=daily, weights=weight)
m <- update(m.adj, . ~ medany + . )
a1<-getAdj(m)[1:2,]
a2<-getAdj(m)[1:4,]
a3<-getAdj(m)[1:5,]
a4<-getAdj(m)[1:5,]

adj<-rbind(a1,a2,a3,a4); adj

# Combine the columns for incidence rates, unadj and adjusted model output and create a CSV file

```
```{r ```
```
```
# Plot
layout(matrix(c(1,2),1),c(1.5,1.05))
# par(mfrow=c(1,2))
par(mar=c(4,4,1,.5))

b2<-t3[c(2:5,1),]
rownames(b2)<-c("0-2","3-6","7-14","15-30",">30")
barplot(b2[,3], ylim=c(0,12), ylab="CDI Incidence", xlab="Time since start of antibiotic therapy (d)", border=NA)

par(mar=c(4,0,1,1))
b1<-t2[c(2:4,1),]
rownames(b1)<-c("0","1-5","6-30",">30")
barplot(b1[,3], ylim=c(0,12), xlab="Time since end of antibiotic therapy (d)", border=NA, axes=F)
```
```{r Sensitivity analysis 1: exclude first Y days}
# Sensitivity analysis 1: >=2, >=5, >=10, >=15, >=20 days of follow up
x<-c("medany","medany_stop_cat","medany_start_cat","medany_sum_cat")
x<-c("medany","medany_stop_5")
y<-c(2,5,10)

for(j in 1:length(y)){
  for(i in 1:length(x)){
    f<-as.formula(paste("~",x[i],"+","bx.t0_hospit",y[j],sep=""))
    m<-update(m.adj,f, data=subset(daily,t0_hospit >= y[j]));
    out[i,j]<-getAdj(m)[2,]
  }
}
```
```
# Sensitivity analysis 2: change of symptom onset date; note that the daily dataset must be reloaded for this analysis
describe(daily$cdi_event2) # N = 130
m<-geeglm(cdi_event2 ~ medany + bx.t0_hospit + as.factor(age4) + male +
               as.factor(n_hospit2) + I(icu_stop<=10) +, id=pid, family="poisson",
corstr="exchangeable",data=daily, weights=weight)
eff.95.geeglm(m,T)[2,]
summary(m)

m<-gee

```
```{r Table 3 part I - antibiotic indexes}
# Incidence Rates
t1<-with(daily,getIRs(cdi_event,medany_stop_5,weight))
t2<-with(daily,getIRs(cdi_event,combi_stop_5_cat,weight))
t3<-with(daily,getIRs(cdi_event,daily$medindex_5,weight))

# Unajdusted
m.un<-glm(cdi_event ~ 1, family=poisson("log"), data=daily, weights=weight)
m<-update(m.un,. ~ medany_stop_5);
u1<-getUn(m)
m<-update(m.un,. ~ as.factor(combi_stop_5_cat) + .);
u2<-getUn(m)
m<-update(m.un,. ~ as.factor(daily$medindex_5) + .)
u3<-getUn(m)

# Adjusted
m.adj<-geeglm(cdi_event ~ bx.t0_hospit + as.factor(age4) + male + as.factor(n_hospit2) + I(icu_stop<=10) + pmin(dp2__10,10), id=pid, family="poisson", corstr="exchangeable",data=daily, weights=weight)
m<-update(m.adj,. ~ medany_stop_5 + .);
a1<-getAdj(m)[1:2,];
m<-update(m.adj,. ~ as.factor(combi_stop_5_cat) + .);
a2<-getAdj(m)[1:3,];
m<-update(m.adj,. ~ as.factor(daily$medindex_5) + .);
a3<-getAdj(m)[1:3,];

# Put it all together
irs<-rbind(t1,t2,t3); irs
un<-rbind(u1,u2,u3); un
adj<-rbind(a1,a2,a3); adj
tab3<-cbind(irs,un,adj); tab3
write.csv(tab3,"Table3a.csv")
print(xtable(tab3),type="html",file="Table3a.html")
```
```{r Table 3 part II - antibiotic classes}
x<-c("med3_stop_5","med19_stop_5","med20_stop_5","med4_stop_5","med22_stop_5","med21_stop_5","med23_stop_5","med8_stop_5","med24_stop_5","med29_stop_5")
y<-vector(length=length(x))
y[1]<-paste("med", c(2,4,5,6,7,8,9,24,26),"_stop_5",sep="","collapse=" + ")
y[2]<-paste("med", c(2,20,4,5,6,7,8,9,24,26),"_stop_5",sep="","collapse=" + ")
y[3]<-paste("med", c(2,19,4,5,6,7,8,9,24,26),"_stop_5",sep="","collapse=" + ")
y[4]<-paste("med", c(2,3,5,6,7,8,9,24,26),"_stop_5",sep="","collapse=" + ")
y[5]<-paste("med", c(2,3,21,23,5,6,7,8,9,24,26),"_stop_5",sep="","collapse=" + ")
y[6]<-paste("med", c(2,3,22,23,5,6,7,8,9,24,26),"_stop_5",sep="","collapse=" + ")
y[7]<-paste("med", c(2,3,22,21,5,6,7,8,9,24,26),"_stop_5",sep="","collapse=" + ")
y[8]<-paste("med", c(2,3,4,5,6,7,9,24,26),"_stop_5",sep="","collapse=" + ")
y[9]<-paste("med", c(2,3,4,5,6,7,8,9,26),"_stop_5",sep="","collapse=" + ")
dliday$med3n_stop_5<-eval(parse(text=paste("with(daily,"y[1],"> 0 )")))
dliday$med19n_stop_5<-eval(parse(text=paste("with(daily,"y[2],"> 0 )")))
dliday$med20n_stop_5<-eval(parse(text="with(daily","y[3],"> 0 )"))
dliday$med4n_stop_5<-eval(parse(text="with(daily","y[4],"> 0 )"))
dliday$med22n_stop_5<-eval(parse(text="with(daily","y[5],"> 0 )"))
dliday$med21n_stop_5<-eval(parse(text="with(daily","y[6],"> 0 )"))
dliday$med23n_stop_5<-eval(parse(text="with(daily","y[7],"> 0 )"))
dliday$med8n_stop_5<-eval(parse(text="with(daily","y[8],"> 0 )"))
dliday$med24n_stop_5<-eval(parse(text="with(daily","y[9],"> 0 )"))`
x <- paste("med", c(3,19,20,4,22,21,23,8,24), "_stop_5", sep="")
z <- paste("med", c(3,19,20,4,22,21,23,8,24), "n_stop_5", sep="")

m.un <- glm(cdi_event ~ 1, family=poisson("log"), data=daily, weights=weight)
m.adj <- geeglm(cdi_event ~ bx.t0_hospit + as.factor(age4) + male + as.factor(n_hospit2)
+ I(icu_stop<=10) + pmin(dp2__10,10), id=pid, family="poisson",
corstr="exchangeable",data=daily, weights=weight)

out <- matrix(NA, nrow=0, ncol=11);
for(i in 1:length(x)) {
  f <- as.formula(paste("cbind(weight,cdi_event) ~",x[i]))
  r <- getIRs.f(f,daily)[2,]

  f <- as.formula(paste("."","-",x[i]," + I(""",x[i],"==0 &",z[i],"==1")," + "."))
  m <- update(m.un,f);
  u <- getUn(m)[2,]

  m <- update(m.adj,f);
  a <- getAdj(m)[2,]

  out <- rbind(out,c(r,u,a))
  rownames(out)[i] <- x[i]
}

rownames(out) <- x;
out

write.csv(out,"Table3b.csv")
print(xtable(out),type="html",file="Table3b.html")
Chapter 4. The Co-Seasonality of Pneumonia and Influenza with *Clostridium difficile* Infection (CDI) in the United States, 1993 to 2008

4.1 Reader’s Note

Few studies have considered the determinants of CDI incidence across a whole country or region. In this manuscript, I demonstrate that increases in hospital pneumonia and influenza precede and predict increases in CDI incidence. A version of this paper was published in the *American Journal of Epidemiology*.


4.2 Abstract

Seasonal variations in the incidence of pneumonia and influenza (P&I) are associated with nosocomial *Clostridium difficile* infection (CDI) incidence, but the reasons why remain unclear. Our objective was to consider the impact of P&I timing and severity on CDI incidence. We conducted a retrospective cohort study using the United States National Hospital Discharge Survey (NHDS) sample. Hospitalized patients with a diagnosis of CDI or P&I between 1993 and 2008 were identified from the NHDS dataset. Poisson regression models of monthly CDI incidence were used to measure (1) the time-lag between the annual P&I prevalence peak and the annual CDI incidence peak and (2) the lagged effect of P&I prevalence on CDI incidence. CDI was identified in 18,465 discharges (8.52 per 1,000 discharges). Peak pneumonia prevalence preceded peak CDI incidence by 9.14 weeks (95% CI: 4.61 to 13.67). A 1% increase in pneumonia prevalence was associated with a cumulative effect of 11.3% over a 6 month lag period (RR = 1.113, 95% CI 1.073, 1.153).
Future research could seek to understand which mediating pathways, including changes in broad-spectrum antibiotic prescribing and hospital crowding, are most responsible for the associated changes in incidence.

4.3 Introduction

*Clostridium difficile*, a toxin-producing bacterium that causes diarrhea, has been identified as the largest single cause of morbidity and mortality among hospital-acquired infections in the Canadian province of Ontario (1). National studies of short-term care hospitals in the United States and Europe show that *Clostridium difficile* infection (CDI) incidence in these jurisdictions has more than doubled since 2000, with increases concentrated in patients over 65 years of age (2–4). Several explanations have been proposed for increasing rates of CDI, and include increasing patient age, acuity, comorbidities, increasing antimicrobial prescribing both in the community and in hospitals and the rapid spread of a previously rare strain (5). Other identified risk factors for hospital acquired CDI include length of stay, exposure to symptomatic cases of CDI (6), and hospital factors including large hospital size and teaching hospital status (3).

Descriptive studies of health care admission patterns demonstrate large wintertime surges of pneumonia and influenza (P&I) hospital admissions, particularly among infants and the elderly (7). Influenza-related admissions are increasingly concentrated in the elderly, and this has been attributed to the increasing frequency of A(H3N2) strain-dominant seasons and the decreasing severity of influenza A(H3N2) seasons for the younger age-groups (8). Changes in CDI distribution demonstrate similar patterns (3). Recently, CDI incidence has been shown to be associated with the seasonal variations in
the incidence of P&I hospitalizations (9), but this may be attributable to increased wintertime hospitalizations rather than increased risk.

If respiratory disease attributable to influenza is an important driver of wintertime surges in CDI, such a link would have important implications for influenza vaccination programs, and could change the risk calculus related to antibiotic prescribing for wintertime respiratory illness. We sought to evaluate the link between the seasonality of P&I and CDI using a large, nationally-representative multi-hospital database from the United States. We sought to do this by (1) estimating the mean annual timing of the CDI peak and the mean annual timing of the P&I peak, and the time-lag between the two (model 1); and (2) ascertaining the magnitude of impact of fluctuations in P&I on subsequent CDI incidence using optimal statistical methods (model 2) (10).

4.4  Methods

4.4.1  Study design and setting

We conducted a retrospective cohort study using the National Hospital Discharge Survey (NHDS) sample. The NHDS consists of diagnosis and demographic data collected from a probability sample of patient discharge records from nonfederal, short-stay hospitals. Due to unequal selection probabilities resulting from the 3-stage cluster sample design, weights are used to make the sample nationally representative. Between 1993 and 2007, approximately 270,000 discharges were sampled annually from a panel of more than 500 hospitals (11). In 2008, the number of sampled discharges was halved. The NHDS has been used previously for the study of CDI incidence (3). The source cohort consisted of all patients ≥40 years of age admitted to an acute care hospital in the United States in the January 1, 1993 to December 31, 2008 period. Younger individuals were excluded due to
low risk of CDI. Patients with length of stay ≤ 2 days were excluded since these patients were at negligible risk of being diagnosed with nosocomial CDI.

4.4.2 Outcome measurement

CDI case patients were identified from the full cohort of hospitalized patients using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) discharge code for CDI (008.45). Compared to toxin-positive test results, the NHDS has good concordance (kappa = 0.64) and neither under nor over-estimates hospital CDI incidence (12). All patients in the NHDS with a primary or secondary diagnosis of CDI, who were not excluded due to young age or ≤ 2 day length of stay, were included. Patients were considered to have acquired CDI in the month of discharge.

4.4.3 Measurement and validation of hospital P&I prevalence

Hospitalized patients with a diagnosis of P&I (ICD-9-CM: 480-488), were also identified from the NHDS dataset. The hospital prevalence of P&I was calculated as the proportion of monthly discharges having a diagnosis of P&I listed as a primary or secondary diagnosis among all admitted patients ≥ 40 years of age. Annual excess P&I hospitalization rates are a strong predictor of excess P&I mortality (8). Also, we validated the correlation between virologic influenza H3N2 strain (%) predominance and maximum hospital P&I prevalence (in the November 1 and March 31 interval) of a given one-year period to ensure that H3N2 was positively associated with hospital P&I burden (13).
4.4.4 Other covariates

We derived clinical and hospital variables including patient age at admission (categorized into 10-year age bins), sex, period of discharge, census region of hospital (Northeast, Midwest, South and West), hospital bedsize, the Charlson comorbidity index and the following 17 specific comorbidities: myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, rheumatologic disease, peptic ulcer disease, mild liver disease, diabetes, hemiplegia or paraplegia, renal disease, malignant cancer, liver disease, metastatic tumor, acquired immunodeficiency syndrome and P&I (14). Discharge month was considered as a continuous variable ranging from 1993 to 2008 by increments of 1/12.

4.4.5 Statistical Analyses

Several time-series analyses were conducted in order to ascertain whether hospital P&I prevalence preceded and predicted CDI incidence. For these analyses, Poisson regression was used; weighted counts per month were modeled with an offset term corresponding to the sum of the weights of all discharges during the month. For each year of data, survey weights \(w\) were rescaled to have a mean of 1 and then divided by the design effect (see Appendix) to obtain rescaled weights \(w^*\), so that monthly case counts equaled the equivalent number of cases under a simple random sample design, also known as the effective sample size (15). We present customary two-sided 95% CIs.

Effect of P&I peak timing

Two harmonic regression models were fitted for both the hospital P&I prevalence and the CDI counts, to determine peak timing and amplitude of average seasonal incidence.
Harmonic regression models are a type of statistical model that explicitly include time as a covariate; they are used to characterize seasonal disease incidence in terms of amplitude (equal to the ratio of peak incidence to trough incidence) and phase shift (16,17) (Figure 3.1). The harmonic regression models were,

\[ \text{N}_{\text{P&I}}(t) \sim \exp \left[ \beta_0 + \beta_1 \sin(2\pi t) + \beta_2 \cos(2\pi t) + \beta_3 t + \beta_4 t^2 + \log(w^*) \right] \]

\[ \text{N}_{\text{CDI}}(t) \sim \exp \left[ \beta_0 + \beta_1 \sin(2\pi t) + \beta_2 \cos(2\pi t) + \beta_3 t (t < 2000) + \beta_4 t (t \geq 2000) + \log(w^*) \right] \]

where \( t \) is month of discharge and January 2000 corresponds to the approximate timing of NAP1 emergence in the United States (18), \( w^* \) is the rescaled weighted discharge count and \( N \) is the rescaled weighted count of cases. For each harmonic model, peak week number (in weeks since January 1st) was obtained by computing \( 52 \tan^{-1} \left\{ \frac{\beta_1}{\beta_2} \right\} \), the inverse tangent of the ratio of the sine function coefficient divided by the cosine function coefficient. The amplitude was equal to \( \exp \left\{ 2(\beta_1^2 + \beta_2^2)^{1/2} \right\} \) where \( 2(\beta_1^2 + \beta_2^2)^{1/2} \) is the amplitude in the log scale. The difference in phase shift parameter between the P&I time-series and the CDI time-series was computed by subtraction. Confidence limits for amplitude, phase-shift and phase shift difference were computed based on the delta method (17).
Figure 4.1: Measurement of a difference in peak timing for two annual harmonic time series. For each time series, the timing of peak incidence or phase shift \((w_1, w_2)\) is calculated using the coefficients of the sine and cosine terms. The phase shift difference \((w_2 - w_1)\) can then be computed by subtraction. On the X axis, week number 0 and 52 correspond to the date of January 1st in consecutive years.

\[\text{Effect of P&I fluctuations}\]

In order to consider the impact of fluctuations of P&I hospitalizations on CDI incidence, we developed a second Poisson regression model, in which CDI incidence was dependent on lagged P&I discharges (10) as well as linear spline trend terms. As such,

\[N_{\text{CDI}}(t) \sim \exp \left[ \beta_0 + \beta_1 q \cdot c_1 + \beta_2 q \cdot c_2 + \beta_3 q \cdot c_3 + \beta_4(t < 2000) + \beta_5(t \geq 2000) + \log(w^*) \right]\]

where \(q\) is a vector of P&I prevalence combining the current and the preceding 23 months, \(c_n\) are polynomial basis vectors and “\(\cdot\)” is the dot product of the two vectors. Specifically,
the polynomial basis vectors were \( c_1 = \{1, 1, \ldots, 1\}, c_2 = \{1,2,3,\ldots, 24\} \) and \( c_3 = \{1,4,9,\ldots, 576\} \).

For all distributed lag models, we removed secular trends from the P&I time-series using a quadratic polynomial, and used the residuals as predictor variables.

*Age-region stratified analyses.* For the harmonic and distributed lag models above, we considered age-region stratified analyses; specifically, we fit models with the same covariates described above to each of the 20 age-region strata, separately. For these analyses, monthly age-region CDI and P&I (rescaled) case counts and weights were extracted. We assessed the pooled effects using inverse-variance weights and a fixed-effects approach (19) and the heterogeneity of study results by use of the \( I^2 \) statistic (20).

In order to consider predictors of age-region specific CDI timing, we created three bivariate linear regression models where age group, region and P&I phase shift were each entered as a covariate for predicting CDI phase shift; measurement error of the P&I phase shift was accounted for using Deming regression.

Analyses were conducted in R using the dlnm and metafor packages (10,21).

### 4.5 Results

Over the 16 year period from January 1st, 1993 to December 31st 2008, there were a total of 4.39 million discharges, of which 1.94 million (44.2%) discharges met the study inclusion criteria (Figure 4.2).
Figure 4.2: Composition of the study population, United States, 1993-2008. All \( n \) correspond to unweighted counts of sampled hospital discharges. NHDS, National Hospital Discharge Survey; LOS, length of stay; CDI, *Clostridium difficile* infection; P&I, pneumonia and influenza.

Over the 16 year period, CDI was identified in 18,465 discharges (8.52 per 1,000 discharges). The incidence of CDI was highest in the Northeastern census region (11.50 per 1,000 discharges) and lowest in the Southern census region (6.84 per 1,000 discharges, relative risk (RR) = 0.59, 95% confidence interval (CI): 0.56, 0.63, Table 4.1). Incidence increased through the four year periods: from 4.70 per 1,000 discharges between 1993 and 1996 to 13.24 per 1,000 discharges between 2005 and 2009 (RR = 2.82, 95% CI: 2.53, 3.13).
Table 4.1: Incidence of CDI (per 1,000 discharges) in Hospitalized Inpatients, United States, 1993 to 2008*

<table>
<thead>
<tr>
<th>Age</th>
<th>Incidence</th>
<th>Relative Risk Estimate</th>
<th>95% CI</th>
<th>Incidence</th>
<th>Relative Risk Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 to 49</td>
<td>4.53</td>
<td>1</td>
<td>Reference</td>
<td>6 to 99</td>
<td>7.31</td>
<td>1</td>
</tr>
<tr>
<td>50 to 59</td>
<td>5.70</td>
<td>1.26</td>
<td>1.10, 1.44</td>
<td>100 to 199</td>
<td>8.52</td>
<td>1.16</td>
</tr>
<tr>
<td>60 to 69</td>
<td>7.52</td>
<td>1.66</td>
<td>1.46, 1.89</td>
<td>200 to 299</td>
<td>10.11</td>
<td>1.38</td>
</tr>
<tr>
<td>70 to 79</td>
<td>9.46</td>
<td>2.09</td>
<td>1.86, 2.35</td>
<td>300 to 499</td>
<td>8.08</td>
<td>1.10</td>
</tr>
<tr>
<td>80+</td>
<td>12.10</td>
<td>2.67</td>
<td>2.38, 3.00</td>
<td>≥ 500</td>
<td>8.96</td>
<td>1.22</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td>Hospital bedsize</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>7.66</td>
<td>1</td>
<td>Reference</td>
<td>6 to 99</td>
<td>7.31</td>
<td>1</td>
</tr>
<tr>
<td>Female</td>
<td>9.18</td>
<td>1.20</td>
<td>1.13, 1.27</td>
<td>100 to 199</td>
<td>8.52</td>
<td>1.16</td>
</tr>
<tr>
<td>Length of stay, d</td>
<td></td>
<td></td>
<td></td>
<td>200 to 299</td>
<td>10.11</td>
<td>1.38</td>
</tr>
<tr>
<td>3 to 5</td>
<td>4.53</td>
<td>1</td>
<td>Reference</td>
<td>≥ 500</td>
<td>8.96</td>
<td>1.22</td>
</tr>
<tr>
<td>6 to 10</td>
<td>8.82</td>
<td>1.94</td>
<td>1.81, 2.10</td>
<td>Reference</td>
<td>6 to 99</td>
<td>7.31</td>
</tr>
<tr>
<td>11 to 20</td>
<td>17.19</td>
<td>3.79</td>
<td>3.51, 4.10</td>
<td>Reference</td>
<td>100 to 199</td>
<td>8.52</td>
</tr>
<tr>
<td>21+</td>
<td>27.88</td>
<td>6.15</td>
<td>5.62, 6.74</td>
<td>Reference</td>
<td>200 to 299</td>
<td>10.11</td>
</tr>
<tr>
<td>Charlon Index</td>
<td></td>
<td></td>
<td></td>
<td>Hospital region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>8.08</td>
<td>1</td>
<td>Reference</td>
<td>Northeast</td>
<td>11.50</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>8.64</td>
<td>1.07</td>
<td>1.00, 1.15</td>
<td>Midwest</td>
<td>8.76</td>
<td>0.76</td>
</tr>
<tr>
<td>2</td>
<td>9.71</td>
<td>1.20</td>
<td>1.11, 1.30</td>
<td>South</td>
<td>6.84</td>
<td>0.59</td>
</tr>
<tr>
<td>3+</td>
<td>8.09</td>
<td>1.00</td>
<td>0.92, 1.09</td>
<td>West</td>
<td>7.60</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Period</td>
<td>4.70</td>
<td>1</td>
</tr>
<tr>
<td>1993 to 1996</td>
<td>11.50</td>
<td>1</td>
<td>Reference</td>
<td>1997 to 2000</td>
<td>5.71</td>
<td>1.22</td>
</tr>
<tr>
<td>1997 to 2000</td>
<td>8.76</td>
<td>0.76</td>
<td>0.72, 0.80</td>
<td>2001 to 2004</td>
<td>9.81</td>
<td>2.09</td>
</tr>
<tr>
<td>2001 to 2004</td>
<td>6.84</td>
<td>0.59</td>
<td>0.56, 0.63</td>
<td>2005 to 2008</td>
<td>13.24</td>
<td>2.82</td>
</tr>
<tr>
<td>2005 to 2008</td>
<td>8.08</td>
<td>1.00</td>
<td>0.92, 1.09</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CDI, clostridium difficile infection.
* For statistical methods relating to bivariate analyses, see Appendix

**4.5.1 CDI Risk Factors**

CDI incidence varied from 4.53 per 1,000 discharges in patients aged 40-49 to 12.10 among patients 80 and over; each ten year increase in patient age was associated with a 28% increase in incidence (95% CI: 1.25, 1.31) of CDI. Female patients were 1.20 (95% CI: 1.13, 1.27) times more likely to have CDI during a hospitalization. Relative to patients with a length of stay (LOS) of 3-5 days, patients with a LOS of 6-10 days were 1.94 times (95% CI: 1.81, 2.10) more likely to be diagnosed with CDI. Charlson index was significantly associated with CDI (F = 7.5, P < 0.001), however the association was not linear. When comorbidities were considered individually we found that pneumonia (RR = 1.50, 95% CI:
1.38, 1.63), heart failure (RR = 1.17, 95% CI: 1.09, 1.26), chronic renal failure (RR = 1.86, 95% CI: 1.67, 2.08), malignant cancer (RR = 1.33, 95% CI: 1.21, 1.45) and acquired immunodeficiency syndrome (RR = 1.90, 95% CI: 1.42, 2.55) were associated with increased risk of CDI while cerebrovascular disease (RR = 0.60, 95% CI: 0.53, 0.69), dementia (RR = 0.74, 95% CI: 0.56, 0.98) and hemiplegia (RR = 0.58, 95% CI: 0.43, 0.78) were associated with decreased risk. The most common comorbid conditions in patients with a CDI diagnosis were urinary tract infections (20.3%), congestive heart failure (19.1%) and pneumonia (10.9%).

4.5.2 Effect of P&I Peak Timing

The Poisson model of monthly CDI incidence included a piecewise linear spline with a change-point in January 2000, in addition to the sine and cosine oscillators (Table 4.2). Peak incidence occurred in week 12 (95% CI: 8.09, 17.02) and was 1.17 times higher than trough incidence (95% CI: 1.07, 1.27). Comparatively, the P&I wave peaked earlier and had a larger amplitude: peak P&I prevalence was 1.63 times higher than trough prevalence (95% CI: 1.56, 1.71) and peaked in week 3 (phase shift = 3.42 weeks, 95% CI: 2.63, 4.21). Peak P&I prevalence preceded peak CDI incidence by 9.14 weeks (95% CI: 4.61, 13.67). The harmonic time-series model for CDI explained 76.3% of variation in CDI. The fitted values of the harmonic prediction model are plotted alongside the overall time series in Figure 4.6 (contained within appendix 4.9.3).
<table>
<thead>
<tr>
<th></th>
<th>Unadjusted</th>
<th>Unadjusted</th>
<th>Adjusted*</th>
<th>Adjusted*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>95% CI</td>
<td>Estimate</td>
<td>95% CI</td>
</tr>
<tr>
<td>P&amp;I prevalence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplitude</td>
<td>1.63</td>
<td>1.56, 1.71</td>
<td>1.59</td>
<td>1.56, 1.63</td>
</tr>
<tr>
<td>Peak (week number)</td>
<td>3.42</td>
<td>2.63, 4.21</td>
<td>3.44</td>
<td>3.03, 3.86</td>
</tr>
<tr>
<td>CDI incidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual Change (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1993-1999</td>
<td>6.8</td>
<td>4.2, 9.5</td>
<td>5.5</td>
<td>3.5, 7.6</td>
</tr>
<tr>
<td>2000-2008</td>
<td>10.8</td>
<td>9.4, 12.1</td>
<td>11.0</td>
<td>9.8, 12.1</td>
</tr>
<tr>
<td>Amplitude</td>
<td>1.17</td>
<td>1.07, 1.27</td>
<td>1.24</td>
<td>1.16, 1.34</td>
</tr>
<tr>
<td>Peak (week number)</td>
<td>12.56</td>
<td>8.09, 17.02</td>
<td>7.70</td>
<td>5.47, 9.94</td>
</tr>
<tr>
<td>Interval between P&amp;I peak and CDI peak (weeks)</td>
<td>9.14</td>
<td>4.61, 13.67</td>
<td>4.26</td>
<td>1.99, 6.53</td>
</tr>
</tbody>
</table>

Abbreviations: P&I, pneumonia and influenza; CDI, *Clostridium difficile* infection; CI, confidence interval.

* Adjusted for census region and age

We applied the same harmonic regression model of CDI to the 20 age-region strata. The pooled estimate of the amplitude was larger than the unstratified estimates (amplitude = 1.24, 95% CI: 1.16, 1.34). No heterogeneity in either amplitude ($I^2 = 0\%$, 95% CI: 0, 44.4, $P = 0.46$) or phase shift ($I^2 = 4.1\%$, 95% CI: 0, 48.0, $P = 0.46$) was detected. When pooled across age-region strata, the CDI peak was estimated to occur in late-February (phase shift = 7.70, 95% CI: 5.47, 9.94). Phase shift of the pneumonia prevalence peak was a significant predictor of CDI phase shift (Figure 4.3, slope = 4.64, 95% CI: 1.31, 7.97), while region ($\chi^2_{3 df}, P = 0.85$) and age ($\chi^2_{4 df}, P = 0.83$) were not. Based on the adjusted differences in phase shift, the *Clostridium difficile* waves for the 20 age-region strata were estimated to peak 4.26 (95% CI: 1.99, 6.53) weeks after the pneumonia waves.
Figure 4.3: Annual timing of CDI peak relative to the pneumonia and influenza peak for 20 age-region strata, United States, 1993-2008. On the X axis, week 0 corresponds to the date of January 1st. CDI, *Clostridium difficile* infection.

4.5.3 Effect of P&I fluctuations

Lag effects in this Poisson model were characterized by a quadratic polynomial crossbasis in order to measure the effect of P&I as a curve across increasing lags. Figure 4.4 presents the impact of P&I prevalence on current and future CDI incidence. A 1% absolute increase in P&I prevalence was associated with a 1.9% relative increase (RR=1.019, 95% CI: 1.012, 1.026) in CDI incidence at lag-month 0. The effect in the months following decreased asymptotically; P&I prevalence had no significant impact on CDI incidence from lag-month 13 onwards. A 1% increase in pneumonia was associated with a cumulative increase in CDI incidence of 11.3% (RR = 1.113, 95% CI 1.073, 1.153) at 6 months and 19.1% (RR = 1.191, 95% CI: 1.084, 1.338) at 24 months. The model explained 78.7% of variation in CDI.
The fitted values of the distributed lag prediction model are plotted alongside the overall time series in Figure 4.5.

Figure 4.4: Distributed lag model estimated effect of hospital P&I prevalence on current and future *Clostridium difficile* incidence, United States, 1993-2008. The unadjusted model effects (squares) represent the lag effects for the single, overall time series, while the adjusted effects (+) are the mean lag effects across all 20 age-region strata. Shaded regions correspond to the 95% confidence bands for the lag effects.

When we applied the same distributed lag model to each of the 20 age-region specific strata, we found an attenuated effect at lag month 0 (RR = 1.009, 95% CI: 1.004, 1.013), and a similar trend of decreasing effect size. The 6-month cumulative effect of a 1% increase in CDI was 4.8% (95% CI: 1.028, 1.067) and 6.2% (95% CI: 1.023, 1.104) at 24 months. In the meta analysis of effect sizes across age-region strata for the 6 month cumulative effect, we found no evidence of heterogeneity ($I^2 = 0\%$, 95% CI: 0, 1.5, $P = 0.95$), although we noted non-significant differences between regions ($\chi^2_{3 df} = 4.07$, $P = 0.25$); lagged P&I effects were smallest in the Midwest (RR = 1.022, 95% CI: 0.994, 1.050) and largest in the South (RR = 1.065, 95% CI: 1.022, 1.110).
4.6 Discussion

This descriptive study has demonstrated that: (1) mean annual peak P&I prevalence precedes peak CDI incidence, (2) hospital P&I prevalence has a dynamic effect on CDI, wherein the effects of increases in P&I on CDI can be felt immediately, and for a period of 12 months, and (3) that these two inferences are also valid within age and region-specific strata. A novel finding of this study is that substantial heterogeneity in the timing of seasonal CDI peak incidence exists for age-region strata and that it was large enough to mask CDI seasonality. A study of CDI seasonality in a smaller, universal, health care system found larger variation in CDI incidence between peak and trough months (22), such healthcare systems may be more susceptible to large-scale CDI epidemics (23) due to greater patient mixing between hospitals (24). Furthermore, our study suggests that that
these variations in peak CDI timing may be associated with the timing of age-region specific P&I peaks. Several studies have found age-specific discharge patterns for P&I, with P&I peaks for younger cohorts earlier than those of working-age cohorts (25). In the United States, P&I discharge peaks in the western states occur 3-4 weeks earlier than in northeastern states (26). In this study, similar patterns were found for CDI.

This study found that increases in hospital P&I prevalence are linearly associated with aggregate CDI incidence. This implies that factors impacting hospital P&I prevalence, including seasonal H3N2 predominance (8) and universal influenza vaccination (27) may impact downstream CDI incidence. Further, this study noted homogeneity of relative risks due to lagged P&I across age and region strata; however, since P&I seasonality is stronger in older age groups, this implies that the risk of infection with Clostridium difficile attributable to P&I is substantially larger in older age groups.

In the only previous study of CDI seasonality (9), which also utilized US medico-administrative data, the authors equally found that CDI peaked in March with CDI tending to follow the number of influenza cases at a lag of 1 month but not at longer lags. In the present study we identified a substantially longer effect of lagged P&I; these findings are likely due to methodologic differences. This study focused on P&I peak timing relative to CDI peak timing using several different statistical methods, including harmonic regression (16), distributed lag models (28), and fixed effects meta-analysis; the effects we identified were robust in the face of varying methodological approaches. Previous studies considering the co-seasonality of infectious disease processes have used regression models with lagged covariates (29) or Box-Jenkins transfer function models with measurement of cross correlation at specific lags (22); in these models a single parameter
or measurement must be made for each lag. Distributed lag models have several advantages over these other models, including reduced collinearity, parsimony, and ease of interpretation.

We hypothesize that the marked CDI seasonality that we observe is likely due to a combination of individual-level factors such as increased uptake of certain antibiotics during wintertime (30) and ecological factors such as increased hospital crowding, CDI pressure (31), herd effects of having many individuals taking more antibiotics and inter-hospital transfer (24). Other system-level studies have shown that effects of P&I and respiratory syncytial virus on CDI are only partly mediated by fluoroquinolone and macrolide prescribing (22). This study provides additional evidence that effects of P&I on CDI incidence exist at the ecological level, and that they are consistent across regions and age groups.

One must note that the mixed ecologic design (32) used in this study, based on data from multiple age and region subgroups, do not represent individual-level associations. Studies using ecologic designs are often critiqued because they are not validated at the individual level (33). Indeed, our exposure of interest, health care system P&I prevalence, was ecologic and as such, individual level inference would have been susceptible to cross-level inference bias (34,35). Instead, we aimed to explore and infer what the aggregate effects that changes in hospital P&I prevalence had on regional CDI incidence. The immediate and persistent effects that increases in P&I prevalence had upon CDI incidence may indicate that both individual-level (i.e. effects on patients with P&I, which should be noted immediately) and herd-level effects (i.e. effects on patients with or without P&I, such as hospital crowding, which may have impacts farther into the future since they act
on the hospital system) are at play. A carefully composed multilevel model, incorporating both individual and group-level measures of P&I exposure, could distinguish the relative contributions of these effects.

Like any observational study, ours has several limitations. First, we did not control for the effects of antibiotic exposure in explaining the association between P&I seasonality and CDI incidence. Indeed, aggregate antibiotic prescribing patterns may partly explain the patterns described in this study. However, since P&I is an upstream determinant of antibiotic prescribing, and since P&I is associated with seasonal respiratory syncytial virus, P&I may be a better predictor for use in early warning systems; in effect, we see intrinsic value in describing the nature of the co-seasonality of P&I and CDI, since the exact causal pathway may not be explained by any single factor (22) and since non-linear dynamic effects may make it difficult to make equivalent inference at the individual level (35). Second, our study considered the incidence of CDI in patients hospitalized for more than 2 days; despite the exclusion of patients hospitalized for ≤ 2 days, our incident cases of CDI were not necessarily hospital-acquired. However, CDI incidence drawn from the NHDS has been shown to accurately measure hospital CDI burden (12), and the effects of P&I on CDI are likely equally large for community-associated disease as they are for nosocomial disease, due to seasonal patterns of outpatient antimicrobial prescribing (36).

In conclusion, this study of CDI over a 16 year period has found that in the United States, seasonal upswings in hospital P&I prevalence precede similar changes in CDI incidence, and that the effects on CDI incidence are long lasting. Proper recognition of CDI peaks, the points in time at which these regularly occur and the impacted groups within
the population could contribute to early recognition of cases and help prevent hospital-system outbreaks.

4.7 Acknowledgments

This work was supported by Mr. Brown’s Frederick Banting and Charles Best Canada Graduate Scholarship from the Canadian Institutes of Health Research (CIHR). Dr. Daneman is supported by a Clinician Scientist Salary Award from CIHR. Dr. Fisman received a grant from the Institute of Population and Public Health of the CIHR.

4.8 References


4.9 Appendix

4.9.1 Bivariate Analyses

For the bivariate analyses (Table 4.1), the incidence of CDI was calculated for each risk group for each NHDS survey year and the standard errors were calculated using the annual, group-specific b parameters provided in the NHDS documentation (11); that is, RSE (p) = sqrt [b × q / (p × SS)], where p is the incidence, q is 1-p, and SS is the weighted total of the number of discharges corresponding to the denominator of p. When group-specific b parameters were not provided for a given risk factor (such as for hospital size groups), we fell back on the overall parameter for the given year. The weighted average of the incidence (weights corresponding to total discharges in each year) and its standard error were then calculated. The standard error of the relative risk (RR) was calculated according to the delta method.

4.9.2 Design Effect

The design effect is a measure of the increase in variance of a given parameter estimate due to sample design. It is equal the ratio of the actual variance of an estimate using a given sample design divided by the variance from a simple random sample with the same sample size (15). Since the variance of the incidence (equal to RSE² multiplied by p²) equals  b × p × q / SS, and the variance of a proportion given a simple random sample design is p × q / ss, where ss is the sample size (unweighted), the design effect is equal to b / (SS / ss) where SS / ss is the mean analysis weight.
4.9.3 Additional Figures

Figure 4.6: Observed CDI incidence and fitted values of harmonic prediction model, United States, 1993-2008. A 3 month central moving average was applied to the observed incidence for this visualization. CDI, *Clostridium difficile* infection. Dotted line: observed incidence. Solid line: predicted incidence.
5.1 Introduction

*Clostridium difficile* infection (CDI) is caused by a unique combination of both intrinsic risk factors that deplete the capacity of persons to resist infection, and environmental exposure to *Clostridium difficile* spores. Recent use of antibiotics, use of medications that reduce stomach acidity, having received a nasogastric tube, having a high comorbidity burden, and receipt of immune suppression medication are all measures that assess individual level risk (1–3). Group-level effects that have been reported in previous studies of CDI are average patient age at the hospital-level, room-level exposure to greater numbers of symptomatic patients, and ICU admission(4–6). The rapport of this thesis to our understanding of *Clostridium difficile* risk factors will be discussed in the following sections.

5.2 Antibiotic Exposure

Antibiotic exposure is the principal risk factor for the development of CDI as it is one of the few mechanisms whereby colonization resistance capacity can be impeded. In two separate chapters of this thesis I explore 1) the effects of specific antibiotic classes and, 2) the duration of antibiotic-associated risks.

While several previous systematic reviews of antibiotic-associated CDI risk focused on inpatient CDI risk, these studies were fraught with study design issues and deemed not appropriate for meta-analysis (7). In the second chapter of this thesis, I performed the first systematic review of community-associated CDI risk, which included a meta-analysis of antibiotic-associated risk. This study contributes to the literature by providing the meta-analysis based estimates of antibiotic class-specific CDI risks. Specifically, I found that:
A. Overall, antibiotic exposure was associated with a 3-fold increase in risk of CDI. Clindamycin, fluoroquinolones, and cephalosporins, monobactams and carbapenems (CMCs) had the strongest associations with CDI risk, while macrolides, sulfonamides and trimethoprim and penicillins had smaller associations with CDI. I noted no association between tetracyclines and CDI risk. In the community setting, there is substantial variation in risk of CDI associated with different antimicrobial classes.

B. Study design had strong impacts on the estimation of CDI-associated effects, and greater standardization of estimates is necessary in future studies. Studies limited to antibiotic-exposed patients demonstrated smaller odds ratios, while studies using a shorter time-window for antibiotic exposure assessment demonstrated larger effects.

One additional weakness of research on antibiotic-associated CDI risk is that studies have not developed a priori predictors of levels of antibiotic risk based on characteristics of the drug, rather than using after-the-fact patient outcomes. Many potential avenues exist for predicting antibiotic-associated risk. Several recent studies have identified the minimum inhibitory concentrations (MIC) of a variety of antibiotics across an impressive number of different Clostridium difficile strains (110 (8) and 606(9)). Antibiotics having the strongest associations with CDI risk in my systematic review appear also to have the highest documented MIC for Clostridium difficile strains in vitro (specifically fluoroquinolones, cephalosporins, and carbapenems) and some of the antibiotics with the lowest MICs had the weakest associations with CDI risk (specifically, tetracyclines). Furthermore, in both these studies the inter-strain variation in MIC for a given antibiotic was lower relative to
the inter-antibiotic variation in MIC. Such systematic characterizations of antibiotics are necessary in order to reliably predict CDI risk associated with antibiotics before their impacts are felt by hospital inpatients.

Most previous studies of inpatient antibiotic-associated CDI risk have considered overall risks associated with any antibiotic exposure on cumulative risk of CDI while hospitalized (10). Although researchers have begun to adopt survival analysis methods for considering the impacts of antibiotic use on risk (11–13), the timing of CDI with respect to antibiotic exposures had not yet been explored. My third dissertation chapter is the first study to provide quantitative estimates of the timing and magnitude of CDI risk due to antibiotic exposures for hospitalized patients. Specifically, I found that:

C. Time since most recent receipt was the strongest independent predictor of CDI incidence. Relative to patients with no prior receipt of antibiotics in the last 30 days, the incidence rate of CDI was 2.41 times higher on days when patients received antibiotics and 2.16 times higher when patients had receipt in the prior 1-5 days. Risk remained elevated 3-6 days and 7-14 days following antibiotic initiation.

D. The adjusted risk associated with CDI was similar when patients either received a single class of antibiotic or when patients received multiple classes of antibiotics. For our antibiotic risk index variable, which was based on established associations of antibiotics with CDI risk, the adjusted risk was elevated, for both exclusive low-risk antibiotics and for high-risk antibiotics.
I provide evidence that antibiotic-associated CDI risks for inpatients are highly elevated for a relatively short period of time after the cessation of antibiotics and then relatively negligible thereafter.

My findings in this study provide empirical measurements of the incubation period for CDI in a hospital context, and of the duration of decreased colonization resistance after antibiotic use. Clinically, this research provides evidence of the importance of preventing unnecessary antibiotic initiation since our findings show that even small doses of antibiotics put patients at substantial risk. In addition, our findings help identify patients at highest current risk of CDI which could aid in more rapid diagnosis, treatment and isolation.

5.3 Age
Advanced age is linked with risk in almost all observational studies of CDI, both for hospital-based and population-based studies. The underlying mechanisms for the association between age and increased risk of CDI may be both intrinsic effects of age on CDI risk and also increased likelihood of being exposed to antibiotics and to virulent *Clostridium difficile* spores in hospital contexts. As such, in population-level studies the effects of age are usually substantially stronger than in hospital-based studies. In one population-level study in the United States, the incidence rates of CDI were almost 8 times higher in the 65 and over population relative to those the 15 to 64 population (14).

In chapter 3, I found that age was linked to CDI risk, even after controlling for disease pressure and antibiotic exposure. Several studies have linked age to differences in gut microflora composition, though these differences are not consistent across regions (15,16). Since the mechanisms and indicators of colonization resistance remain poorly
understood, studies linking age-related differences in gut microflora function shed little light on the reasons for strong age-related effects. A more likely explanation is that increasing age is associated with immune response to vegetative Clostridium difficile bacteria and Clostridium difficile toxins. Indeed, patients that become colonized with Clostridium difficile and are also able to mount an immune response to Clostridium difficile and its toxins are less likely to have overt symptoms (17). Also, patients that become colonized with toxigenic Clostridium difficile and develop high levels of anti-toxin IgG are less likely to develop diarrhea (18).

In chapter 4, I found that hospital pneumonia and influenza (P & I) prevalence is a strong predictor of hospital CDI risk. Indeed, mean inpatient age is correlated with seasonal P & I dynamics (19). I found that P & I effects on CDI risk are smaller but still substantial in age-region subgroups. This suggests that part of the reason for large P & I effects on CDI risk are mediated by the impacts of P & I on inpatient age structure. Thus, the increasing frequency of H3N2 influenza seasons (20), which disproportionately impact the elderly and bring them into hospital, may be contributing to increasing CDI incidence.

Most hospital-based studies of CDI have ignored the distinction between the direct and indirect effects of age on CDI. One recent study of CDI across 64 hospitals in the United States demonstrated that mean hospital inpatient age was associated with CDI risk, even after controlling for age at the individual level (4).

5.4 Infection Pressure

Two previous studies have considered Clostridium difficile disease pressure as a risk factor for incident CDI. In those studies, disease pressure was measured as the number of patient-days of exposure to patients with CDI at the ward level (21,22). In both studies,
patients were considered at risk of spreading disease for 14 days from positive test date. Such CDI pressure measures are a rough proxy of spore exposure levels for susceptible patients in hospitals.

E. In chapter 3 of my dissertation, I demonstrated that there was no association between *Clostridium difficile* disease pressure from identified symptomatic patients and inpatient CDI risk.

One possible reason is that symptomatic CDI cases are not actually responsible for disease spread at Sunnybrook hospital. Recent research on CDI transmission using whole-genome sequencing in order to identify transmission chains suggests that disease pressure effects from symptomatic patients may be small. Specifically, a putative link between identified CDI could not be found for 45% of CDI cases (23). Further, the study showed that the number of CDI cases with identifiable hospital links decreased from over 10 cases per month to less than 2 cases per month over the 4-year study period. Meanwhile the incidence of cases with no identifiable link remained relatively stable at 10 cases per month (Figure 5.1). As such, one potential explanation for our null results is that the effects of disease pressure were below recognizable levels due to adequate infection control practices and effective treatment at Sunnybrook hospital. Furthermore, in a hospital setting where the incidence of CDI is constant through time, as was the case in our cohort, the impact of disease pressure is likely to be very small, since otherwise, the dynamic nature of the infectious disease would drive rates higher.

Thus, my research suggests that sources of infection other than identified infectious patients, such as asymptptomatically colonized patients, may be responsible for the spread of disease. Further study considering colonization pressure from various types of
potentially infectious patients must be considered. Knowledge of sources of infection is essential in order to be able to practice evidence-based infection control.

Figure 5.1: Monthly case counts at a hospital in Oxfordshire, England (23)

5.5 Pneumonia and Influenza Effects and Seasonality

Although previous research suggested an association between the time-trends of pneumonia and influenza and CDI case counts, in chapter 4 of this thesis, I provided evidence that actual risk, not just increased wintertime hospitalization rates, explains the observed seasonality. Further, this is the first time the duration of increased risk due to P & I hospitalizations has been quantified.

F. I demonstrated that: 1) mean annual peak P&I prevalence precedes peak CDI incidence, 2) hospital P&I prevalence has a dynamic effect on CDI, wherein the effects of increases in P&I on CDI can be felt immediately, and for a period of 12 months, and 3) that these two inferences are also valid within age and region-specific strata.
G. Secondary findings of this thesis are that a substantial heterogeneity in the timing of seasonal CDI peak incidence exists for age-region subgroups and that it was large enough to mask CDI seasonality.

The long duration of estimated P & I effects (13 months), in addition to the disproportionately large effects (a 1% absolute increase in P & I, which translates to an approximate 3% relative increase, was associated with a 13% relative increase in CDI incidence) means that small season-to-season differences in influenza dynamics can have incredibly large effects on CDI risk. During the 16-year study period, substantial secular trends in CDI incidence undergirded the seasonal trends we observed. My models suggest that at least part of those secular trends were explained by P & I. Specifically, consecutive severe P & I seasons leading to greater hospitalization could lead to rapid CDI incidence increases, while consecutive low P & I seasons could lead to multi-annual decreases in CDI incidence.

The estimated risks I presented represent the ecological effects of P & I prevalence on CDI incidence, and as such must be interpreted at the ecological level. The causal mechanisms underlying the observed association between P & I is increased mean patient age, increased emergency department and hospital crowding (24), and increased antibiotic prescribing (25) (Figure 5.2).
Figure 5.2: Suggested pathways for the association between seasonal respiratory infections and CDI risk.

However, increases of patient and hospital-level predictors do not directly account for their disproportionate impact of P & I on CDI risk. Mathematical modeling approaches can shed further light on the large observed seasonal fluctuations in CDI incidence. Seasonal forcing due to a change in transmissibility of an infection can result in such disproportionate impacts on observed incidence (26). Soper argued that changes in the average monthly contact coefficient for measles was responsible for annual measles periodicity (27). As Keeling later demonstrates using a mathematic model of measles, a 10% wintertime increase in the transmission parameter may translate into a 78% increase in case notifications (26). Using a simple compartmental model of hospital CDI, I demonstrated that a 13% increase in wintertime transmission rates coinciding with the pneumonia and influenza seasonality could account for the 50% increase in springtime CDI incidence as observed in Quebec (28).

The use of multilevel models explicitly incorporating characteristics of influenza seasons (29), could help characterize and preempt years when CDI incidence will increase the most. Such modeling could help develop predictive models that are commonly used by
public health officials and help to plan the allocation of public health resources (30). A CDI prediction model could help preempt increases in CDI incidence and aid health care systems cope with CDI emergence. Assessment of the utility of different prediction time-scales (1, 3, or 6 months) would have to be determined, however.

Seasonal changes in hospital admissions are strongly associated with CDI incidence at a hospital-network level. Interventions that could reduce the effects of seasonal P & I on the hospital network could also reduce CDI incidence. Namely, universal influenza vaccination, programs to reduce hospital admissions among elderly persons with community-acquired pneumonia, and antimicrobial stewardship programs could potentially reduce CDI hospital network CDI rates.

5.6 Innovative Uses of Statistical Methods

Multiple methods developments were made in the analyses I conducted for this thesis. I will briefly discuss these innovations as they apply to chapters 2, 3 and 4 of this thesis.

5.6.1 Meta-Regression of CDI Risk Factors

In the meta-analysis of antibiotic-associated risk factors for community associated CDI, I used meta-analysis techniques in order to better understand inter-antibiotic class heterogeneity in effects of antibiotic risk. Previous studies stated that low study quality made meta-analysis unfeasible (7). My approach was to use Anatomical Therapeutic Chemical classes (31) in order to create a standardized categorization of antibiotics. When many different antibiotics of the same class were reported in a given study, these were collapsed into a single effect size. This allowed for a relatively granular analysis of antibiotic associated risks. This approach has since been used for a meta-analysis of
antibiotic-associated risks that included both hospital- and community-associated CDI cases (32).

### 5.6.2 Cohort Studies of Hospital Acquired Infections

Most statistical methods development in epidemiology has focused on chronic diseases such as cancer and cardiovascular disease that develop slowly over a long period of time, and are not driven by contagion (33). As such, some of these methods may not be optimal for understanding nosocomial infections that develop rapidly over the course of days and hours from the time of hospital admission, and are driven by contagion.

Rather than use logistic regression based on cumulative exposures over the course of a hospitalized patients’ risk period, my analyses were based on a Poisson time-to-event model and incorporated both: 1) weighted cumulative exposures to account for the time-sensitive nature of exposures (34), and 2) disease pressure, to incorporate the impacts of contagion between CDI cases. Further, analyses of contagious diseases should always attempt to incorporate risk factors related to the sources of transmission, as exposure to contagious particles is an intrinsic risk factor in any contagious disease outcome and ignoring contagion may result in the incorrect estimation of the impacts of risk factors.

Although data for these types of analyses are routinely collected within hospital pharmacy, patient location, and microbiology laboratory databases, the data structure and methods for analyzing time-varying exposures are difficult to implement compared to logistic regression methods. I developed a statistical analysis strategy that enabled the efficient analysis of hospital infection control data involving both lagged time-varying exposures and measures quantifying disease contagion. Four principal steps are involved in this strategy.
First, the information on at risk patients must be transformed from their native format in hospital databases into what is known as counting process format (35) wherein patient characteristics and exposures are recorded in a table for every patient-day of hospitalization. Second, exposure and contact patterns between the susceptible cohort of interest and potentially infectious and symptomatic patients must be extracted from hospital patient location databases (22,36). Third, for each patient-day, exposure history for antibiotics, contact patterns, and other treatments must be lagged from previous days using a weighted function of previous days’ exposures (34). Fourth, since the size of these datasets becomes quite large (since there are records for each patient-day), efficient data handling is needed. I developed a series of R functions that allow the efficient modeling of this structure of data (R library forthcoming). The two most distinctive features of this analysis strategy are that: 1) both at-risk patients and infectious patients are retained in the principal analysis dataset, and that 2) functions are used to lag risk factors and calculate disease pressure on the fly, rather than being stored and kept in physical datasets. In addition to being applicable to understanding Clostridium difficile risk factors, this analysis strategy could be adapted to model other hospital-acquired infections.

5.6.3 Co-Seasonality of CDI and P & I

In this study of the seasonal variations in CDI risk across the United States, I used a distributed lag model to assess the impacts of P & I on CDI risk. In particular, I used a distributed lag model in order to assess the combined impacts of multiple previous months’ hospital-level P& I burden on current CDI risk. As I have discussed in other sections of this thesis, these methods have been used to study the impacts of weather patterns on all-cause and cause-specific mortality, but had never been used to study
disease seasonality in the past. This approach allowed us to show that the impact of P & I on CDI incidence in the US health care system extends for a long period after P & I itself has disappeared and suggests that the impacts of P & I on CDI risk may be due to ecological risk factors such as increased contagion due to increased levels of antibiotic use.

5.7 Limitations

Although I have highlighted manuscript-specific limitations in chapters 2 to 4, here I aim to highlight several of the limitations that cut across all manuscripts of this thesis. In particular, these limitations were associated with the sensitivity and specificity of the CDI diagnosis, the distinction between community- and hospital-associated disease, and the ascertainment of antibiotic exposures.

CDI diagnosis involves clinician judgment with respect to when a Clostridium difficile diagnostic test should be performed; this threshold is imprecise and may have varied between studies (Chapter 2), between wards (Chapter 3), or between hospitals (Chapter 4). As discussed in detail below (section 5.9.2), CDI testing rates themselves may partially drive international differences in CDI incidence rates, with higher-GDP countries that test more frequently reporting higher rates of CDI than countries that test less frequently. Further, study 1 was based on clinical diagnoses of CDI across many different hospitals and as such would have been subject to different diagnostic criteria and testing methods. Indeed, PCR methods are more likely to detect Clostridium difficile toxin and this has been shown to be associated with increased reported CDI incidence rates and decreased severity of detected cases (37). To a lesser extent, the CDI testing threshold may have differed across wards or individual health care workers in Chapter 3.
In addition to issues of related to CDI diagnosis, both studies 1 and 2 were potentially impacted by the attribution of the origin of infection as being community-associated or hospital-associated. Very little research has compared the etiologies and risk factors for these subgroups of CDI cases. In chapter 2, 2 of the component studies may have misleadingly been studying CDI of hospital-origin; it is unclear how this may have affected the results. In chapter 3, only patients that developed first-time CDI infections of hospital origin as determined by Infection Prevention and Control staff were included in the case group. Thus, the study excluded all patients that developed disease within 48h of arrival at the hospital, some of whom may have had new-onset CDI caused principally by hospital spore exposures. Since differences in risk factors for community- versus hospital-acquired CDI are not well known, a future meta-analysis of antibiotic-associated CDI risk may wish to compare and contrast risk for CDI of community- and hospital-origin. This could augment the small number of studies of antibiotic associated risk among outpatients [32], and help determine whether the antibiotic-associated risk profiles for inpatient and outpatient sub-populations are distinct.

The assessment of antibiotic exposures was challenging in each of the studies. In chapter 2, the component studies of the systematic review assessed antibiotic exposure time-windows varying from 42 to 120 days prior to the CDI onset. Heterogeneity of the antibiotic-specific effect sizes between studies may have been partially due to these differing time-windows (shorter time-windows appeared to be associated with larger relative risks). In chapter 3, there was no means of identifying patient antibiotic exposures prior to hospital admission in the described cohort study. As a sensitivity analysis, I measured the impacts of antibiotic exposures in a restricted cohort composed of all patients that had been admitted for at least 10 days prior to symptom onset and found no
differences in the measured effect sizes. Nevertheless, this lack of antibiotic prescribing information prior to hospitalization was an important limitation because impacts of outpatient antibiotic use could persist into the period after admission. Finally, in chapter 4, I had no measures of national inpatient or outpatient antibiotic exposures. Having this information would have allowed for a mediation analysis (38) to better understand what proportion of the P & I effect on CDI risk was associated, or mediated, by increased antibiotic exposures.

5.8 Implications

The principal findings of this thesis, that, antibiotic utilization itself, and P & I seasonality (presumably mediated by an upsurge in antibiotic prescribing) are important drivers of CDI risk, suggest avenues for clinical and public health practice. Specifically, antimicrobial stewardship programs that aim to reduce inappropriate antibiotic utilization, and influenza immunization and other programs that reduce the impacts of P & I on the health care system, could help reduce CDI incidence in North America.

With respect to antibiotic utilization, this thesis has both clinical and population-health implications. Clinically this thesis suggests that physician prescribing practices should aim to reduce the inappropriate initiation of antibiotic therapy and reduce the overall duration of therapy in order to keep patient risk at a minimum. However, in order to achieve transparent and quantifiable improvements in prescribing practices, institutional and regional antimicrobial stewardship programs should guide these efforts. Hospital antimicrobial stewardship programs consistently achieve substantial reductions (22%-36%) in overall antibiotic utilization and as such are generally seen as cost-effective (39). Since a large portion of CDI cases occur without patients having been admitted to a
hospital (40), comprehensive regional antimicrobial stewardship programs that include oversight of primary care prescribing should also be implemented (41).

In addition to antimicrobial stewardship, this thesis suggests that interventions aiming to reduce the burden of P & I on the health care system would be effective at reducing CDI incidence. Implementation of universal influenza immunization in October 2000 was shown to reduce influenza-associated hospitalizations by 42% in Ontario, relative to other Canadian provinces without universal influenza immunization programs (42). This program resulted in, as an unintended effect, a large reduction in the prescribing of respiratory antibiotics (43). In addition to universal influenza immunization, clinical pathways for treating long-term care facility residents with respiratory infections on-site could also reduce the impact of P & I on the hospital system (44). This thesis suggests that such interventions could help stem the rising incidence of CDI in North America.

5.9 Potential Areas for Further Research and Methods Development

5.9.1 Clostridium difficile Transmission Sources

Since a principal finding of this thesis was that increased levels of Clostridium difficile disease pressure were not associated with increased CDI risks, further study is needed to identify the actual sources of CDI spore exposure. In addition to being risk factors for overt disease, patients with Clostridium difficile risk factors (such as increased age, antibiotic use and gastric-acid suppressor use) are at increased risk of developing asymptomatic disease (18). As such, important risk factors for CDI disease may also impact risk by increasing risk of transmission (45). It would be pertinent to conduct a cohort study wherein new metrics quantifying patient-days of ward-level exposure to patients at risk of asymptomatic colonization (such as antibiotic users) were considered as risk factors for
CDI outcomes. Ward-level antibiotic exposures can also be used as proxy measures of potential contact with *Clostridium difficile* colonized patients. Depending on the scale of the project, this study could be supplemented with measures of environmental contamination in order to help show that increased exposure to contaminated environments acts as the intervening pathway. Additionally, routine measures of patient colonization status would be helpful, but may not be financially feasible at scale that would allow detection of an association with CDI outcomes.

### 5.9.2 Impacts of CDI Testing Rates on CDI Incidence Measurement

Inter-institution comparisons of infection rates rely on infection endpoints that accurately reflect true incidence of disease and that are consistently measured across centers. A recent article by Haley and colleagues takes important steps towards improving the reporting on CDI rates (46). In their study of 3,458 reported hospital onset CDI cases in 124 hospitals in New York state, they assess the potential for three measures: numerator audit, denominator correction, and age-adjustment, to improve accuracy of hospital incidence classification. Combined, these three measures do not result in much reclassification: 6% of hospitals are reclassified into higher risk groups and 6% are reclassified into a lower risk group. Further, the most influential of the 3 factors was denominator correction, and this correction is easy to implement: hospitals only need to use their information systems to subtract hospitalized-days of patient-stays of less than 4 calendar days. All in all, it’s an easy message to relay to hospital systems instituting mandatory reporting of CDI rates: “mind your denominator!” (47) But should we really be consoled, or are there other issues with CDI reporting lurking below the surface?
One potential source of bias that has not been addressed in the literature on CDI reporting is CDI testing rates. Figure 5.3 panel A shows a 14-fold variation in *Clostridium difficile* testing rates (from <10 to 140 tests per 10,000 patient days) across tertiary hospitals in European countries, that correlates strongly with CDI incidence ($R^2=0.64$, data retrieved from (48)). This relationship may in fact reflect the higher incidence of CDI in high-testing countries, since increased test positivity may spur increases in testing levels (49), however, as Figure 5.3 panel B shows, there is no such correlation ($R^2=0.00$).
Figure 5.3: Panel A. The association between *Clostridium difficile* testing rates and, in Panel A, reported *CDI* rates, in Panel B, the per-test positivity of *Clostridium difficile* test results. Data are from a secondary analysis of a study of 97 hospitals across 34 European countries and both regression lines are based on simple linear regression of the 28 countries providing testing, patient follow-up time, and incidence data (48).

The National Healthcare Safety Network (NHSN) surveillance definitions attempt to standardize testing rates (50). Specifically, all unformed stool specimens that are sent to the hospital laboratory are subjected to CDI testing, and repeat specimens within 2 weeks are considered duplicates and not reported; but these measures do not specify who should, and who should not receive a test. Are all patients with diarrhea tested, or only a portion? And, what is considered diarrhea (51)? These ambiguities suggest that the symptom severity threshold for initiating testing could vary significantly between institutions and wards. In addition, use of more sensitive PCR-based tests may result in a substantial reporting of low-severity CDI cases and *Clostridium difficile* carriers who develop diarrhea for another reason (37).

Until we take measures to quantify and understand the relationship between reported incidence, frequency and mode of testing, and the frequency of complicated *Clostridium difficile* infections, reported CDI rates could depend on how hard hospitals look for what they don’t want to find.

### 5.9.3 Prediction of CDI Spread in Health Care Networks

An extension of the time-series methods used for modeling CDI seasonality could be applied to study the emergence of CDI in the Ontario health care system. A study of the California health care system showed that institutions are highly and heterogeneously interconnected by patient sharing (52). This study quantified two interesting features of hospital networks. First, they found that direct transfers only accounted for a fraction of total hospital connections whereas connections with intervening nonhospital stays were substantially more prominent. Second, they quantified hospital in-degree and out-degree, representing tendencies of a given hospital to be influenced by other hospitals in the
network (via receipt of patients) or to exert influence on other hospitals (via sending of patients): hospitals with high in-degree tended to be smaller, whereas hospitals with high out-degree tended to be larger.

Other studies (53,54) have explored how inter-hospital patient transfer may play a role in the spread of MRSA, using extensive simulations based on the same Orange County network data. One study (53) revealed that outbreaks in larger hospitals tended to lead to larger sustained increases in MRSA incidence across hospitals in the network, while smaller hospitals tended to be more influenced by outbreaks in other network hospitals (see Figure 5.4). As such, studies have shown how to measure hospital interconnectedness, potential for inter-hospital transmission, and the expected size and duration of network-wide effects engendered by hospital outbreaks in a given hospital.

Figure 5.4: The effect of outbreaks on network and the effect of outbreaks on hospitals, according to hospital size (53)

However, no studies to date have joined data on patient sharing and disease incidence in order to identify whether and to what extent patient sharing actually translates into real changes in disease incidence. Thus, future studies could ascertain the impact of patient sharing on real inter-hospital *Clostridium difficile* dynamics, in order to:
• Develop a statistical model to predict changes in the weekly hospital incidence of CDI in a health care system, based largely on our previous work on *Clostridium difficile* in the United States (55);

• Measure “external” disease pressure which is disease pressure emanating from CDI-infected patients that have recently visited another hospital;

• Estimate the effects that transfer of patients infected with CDI and patients emanating from high CDI incidence hospitals have on CDI incidence in low-incidence hospitals.

Due to the availability of a comprehensive provincial dataset of hospital admissions, and the reliable coding of CDI in medico-administrative databases, Ontario would represent an ideal laboratory for understanding inter-hospital dynamics of CDI.

Two different measures of potential for inter-hospital transmission, measured weekly, could be used to quantify the impact of external CDI incidence and prevalence on CDI incidence in a given hospital. The first measure could be the number of patient-days contributed by patients having a previously recorded diagnosis of CDI that was first recorded at another hospital. The second measure could be the number of patients received from another hospital (whether via uninterrupted or interrupted patient sharing) weighted by the prevalence of CDI in the other hospital at the time of the patients’ most recent stay (weighted external prevalence). Such a study would be one of the first to model the impacts of patient movement on actual CDI spread. Similar methods have been used successfully for modeling the spread of cholera in Haiti (56).
5.9.4 Statistical Modeling Strategies and the Independent Censoring Assumption

One limitation of all studies based on hospital inpatients is that little is known about what happens to patients after discharge. I overcame this problem by considering discharged patients as censored. However, standard censoring methods assume that patients are censored at random, or more specifically, that the probability of censoring is not related to the outcome (in this case CDI). Inverse probability of censoring models have been developed to overcome this limitation (57,58), but these have rarely been used in hospital infections research (59). One avenue for future research is to explore the impact and interpretation of different modeling strategies. The three most obvious options are 1) logistic regression, 2) survival analysis with the assumption of independent censoring, or 3) survival analysis with inverse probability of censoring weights. A systematic comparison of these models may shed light on the most appropriate and statistically efficient methods for the estimating CDI risk factors.

5.10 Conclusion

This thesis considered individual and ecological risk factors for CDI. I have provided evidence that antibiotics, with the exception of tetracyclines, are consistently associated with CDI risk across many studies but that the effects of antibiotics are relatively short-lived, with risks, in a hospital context, concentrated during and within 5 days of antibiotic cessation. I also showed that the impacts of seasonal pneumonia and influenza dynamics on hospital-network CDI dynamics were substantial, and the effects of a severe P & I season lasted for 13 months. As such, this research suggests that controlling the Clostridium difficile epidemic requires a coordinated response that is implemented across an entire hospital network, and that this response should involve an appreciation of the
dynamic nature of *Clostridium difficile* disease. Measures that lessen the impact of P & I on the health care system, including universal influenza vaccination, programs to prevent hospital admissions of elderly persons with community-acquired pneumonia, and antibiotic stewardship programs to reduce both overall antibiotic use and wintertime surges in antibiotic prescribing are all avenues that my thesis research indicates would be effective in lessening the current North American CDI epidemic.

5.11 Citations


