2022-02-02

On the Optimization of Clostridioides difficile Diagnostics Through RT-PCR Cycle Threshold Defined Zones of Disease Probability

Doolan, Cody Patrick

UNIVERSITY OF CALGARY

On the Optimization of *Clostridioides difficile* Diagnostics Through RT-PCR Cycle Threshold Defined Zones of Disease Probability

by

Cody Patrick Doolan

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN MICROBIOLOGY AND INFECTIOUS DISEASES

CALGARY, ALBERTA

FEBRUARY, 2022

Abstract

On the Optimization of *Clostridioides difficile* Diagnostics Through RT-PCR Cycle Threshold Defined Zones of Disease Probability

Advisor

Dr Dylan Pillai

Cody Patrick Doolan

University of Calgary 2022

Clostridioides difficile is an opportunistic pathogen with a large burden of disease and no gold standard test. Quantitative polymerase chain reaction (qPCR) offers excellent sensitivity but overcalls clinical *C. difficile* infections (CDI) due to high prevalence of colonization. The hypothesis of this thesis is that the CDI qPCR results can be titrated down to determine clinical CDI more accurately and possibly aid in predicting disease severity.

A cross sectional study was conducted on suspected CDI patients evaluating if qPCR cycle threshold (Ct) can be correlated to probability of CDI. Latent class analysis (LCA) was employed with observed variables including four commercial qPCR tests, toxin detection by enzyme immunoassay, toxigenic culture, fecal calprotectin, and CDI clinical diagnosis. Three defined zones as a function of qPCR cycle threshold (Ct) were identified: CDI likely (>90% probability), CDI equivocal (<90% and >10%), CDI unlikely (<10%). A single model comprising toxigenic culture, clinical diagnosis, and toxin EIA demonstrated the best fitness.

The following Ct cut-offs for 4 commercial test platforms were obtained to delineate CDI probability zones: GeneXpert®: 24.00, 33.61; Simplexa®: 28.97, 36.85; Elite MGB®: 30.18, 37.43; and BD Max™: 27.60, 34.26.

A prospective cohort study was conducted to investigate if these zones (using GeneXpert®) can be further correlated to indicators of severe CDI. Primary diagnosis, demographic data and known indicators of disease severity were captured: white blood cell, creatinine, albumin, C-reactive protein, and hospital length of stay. A sub analysis was conducted evaluating a subset of the patient population attempting to isolate patients whose clinical variables were most influenced by CDI. No significant correlations were found between the clinical variables investigated and Ct values or Ct zones.

This work establishes a method of using currently deployed diagnostics to allow clinicians to reduce overdiagnosis of CDI. Decreasing false positives could have broad impacts, increase targeted treatments, and decrease antibiotics used with mistreatment. The average cost attributed to CDI for one patient is estimated at $11,917. LCA models predict that qPCR confirmation overdiagnoses patients in Calgary by at least 20.9%. If CDI confirmation were reduced by 20.9% this could equate to massive savings; Foothills Medical Center alone could save over $929,000 annually with no additional investment in laboratory infrastructure.

Preface

Chapter 2. Is published in Clinical Infectious Disease and is used with permission of Doolan et al. ©The Author(s) 2020 (See Chapter 6.0 Appendix).

Chapter 3. Is used in this thesis with permission from the author.

Table of Contents

List of Figures

List of Abbreviations

CDI	*Clostridioides difficile* infection
RT-PCR	Real Time Polymerase Chain Reaction
GS	Gold Standard
LCA	Latent Class Analysis
CCFA	Cycloserine Cefoxitin Fructose Agar
NAAT	Nucleic Acid Amplification Test
PCR	Polymerase Chain Reaction
CCCNA	Cell Culture Cytotoxicity Neutralization Assay
EIA	Enzyme Immunoassays
GDH	Glutamate Dehydrogenase
Ct	Cycle Threshold
ROC	Receiver Operator Curves
AUC	Area Under the Curve
PCR	Polymerase Chain Reaction
tcdA	*C. difficile* Toxin A gene
tcdB	*C. difficile* Toxin B gene
tcdC	Binary toxin gene marker for ribotype 027/NAP1 strain
CCCNA	Cell Culture Cytotoxicity Neutralization Assay
GDH	Glutamate Dehydrogenase
IDSA	Infectious Disease Society of America
fCPT	Fecal Calprotectin
LOD	Limit of Detection

BIC	Bayesian Information Criterion
EIA	Enzyme Immunoassays
FMC	Foothills Medical Center
APL	Alberta Precision Laboratories
ICU	Intensive Care Unit
EMR	Electronic Medical Records

1.0. Introduction

1.1. Background

Clostridioides difficile (previously *Clostridium difficile*) is commonly accepted as the leading infectious cause of nosocomial diarrhea in developed countries [1-3]. It is the organism responsible for almost all cases of pseudomembranous colitis, 10-25% of antibiotic associated diarrhea, and presents a large burden of morbidity and mortality [4]. In 2011, in the United States alone, this bacterium caused 500,000 infections and 29,000 deaths with an annual cost in excess of $4 billion [5]. Canadian incidence estimates range from 34,139 to 41,725 cases annually [6].

To understand the intricacies of clinical diagnostics of *C. difficile* infection (CDI), it is crucial to understand the organism and pathology of CDI. *C. difficile* is an anaerobic, spore forming, Gram-positive, toxin-producing bacillus with many environmental sources (livestock & wild animals, seafood, vegetables, healthcare, human carriers, etc.) [4, 7, 8]. This bacilli was first described in 1935 but was not appropriately linked with antibiotic-associated diarrhea until the 1970s [3]. The spores of *C. difficile* can persist in the environment for over 12 months resulting in patient to patient transmission and spread from environmental fomites [3]. Contaminated hospital surfaces play an important role in the transmission of *C. difficile* as spores are resistant to standard cleaning compounds [9].

In-hospital colonization of patients is quite common with reported values as high as 20-30% of inpatients carrying the organism [3, 8]. It is generally agreed upon that there is no need to treat asymptomatic carriage of toxigenic *C. difficile,* but these patients do play a role in introducing and maintaining CDI transmission in hospitals and communities [4, 8]. As not all patients

colonized with *C. difficile* develop CDI, it is suggested that possible immune response and microbiota balance are important in the pathogenesis [4]. Infants and young children are common carriers of *C. difficile* in their gut microbiome and much of the time appear to have no symptoms even when colonized with a toxigenic strain [3]. While hospital-acquired CDI is the most common, it is important to note there is an increasing prevalence and burden of community-acquired CDI and the application of the findings of this thesis will not be limited to nosocomial infections [8].

Early detection of disease and isolation of CDI patients, hand hygiene, antibiotic stewardship, and environmental cleaning remain important preventative measures [9]. The primary risk factor in development of CDI for both hospital and community acquired infections is a disruption of the gut microbiome caused by antibiotics [3]. The onset of symptoms generally occurs within 2 weeks of initiation of an antibiotic treatment with fluoroquinolones, broad-spectrum penicillin, broad-spectrum cephalosporins, and clindamycin being frequent culprits [8]. Other common risk factors for CDI include advanced age (>65 years), immunosuppression, inflammatory bowel diseases, cancer chemotherapy, and exposure to healthcare setting [5]. Proton pump inhibitors have additionally been found to increase risk of CDI [5, 9].

CDI pathophysiology involves the colon where *C. difficile* spores transition to a vegetative form and produce two major exotoxins, toxin A (TcdA), and toxin B (TcdB) [4, 5]. Toxin A and B interact with the colonic epithelium disrupting tight junctions and breaking down the actin cytoskeleton of enterocytes, while causing cytokine/chemokine production, increased neutrophil infiltration, increased enterocyte cell death, and disruption of the intestinal mucosal barrier [5].

Clinical manifestation ranges from asymptomatic carriage, to diarrhea, abdominal cramping and/or distension, fever, leukocytosis, colitis with and without pseudo membranes, and can progress to toxin megacolon, perforation, sepsis, and death [3]. There is no consensus on a scoring method for categorizing disease severity or disease status. Disease severity as suggested by the Infectious Disease Society of America (IDSA) and used by the Association of Medical Microbiology and Infectious Diseases Canada (AMMI), classifies severe CDI disease when leukocytes are ≥15,000 cells/ml and/or serum creatine is ≥1.5 mg/dl [10, 11]. Other disease scoring systems do exist with varying predictive value. With a lack of a clear and primary disease classification system clinical judgement remains the primary and most widespread method to evaluate and manage individual patients [10].

There are several hypervirulent strains of *C. difficile* associated with increased incidence and severity of CDI [5]. The most common hypervirulent strain is NAP1/027. NAP1 is characterized by high sporulation, high toxin production, increased treatment failure, high mortality, and fluoroquinolone resistance [5].

1.2 Laboratory Diagnostics

In the late 1990's, CDI surveillance and detection methods dramatically improved with a corresponding increase in the incidence and severity of observed CDI [12]. This is due not only to outbreaks of hypervirulent strains with increased transmission, but also increases in our ability to detect both the organism and disease process [12]. How to best utilize highly sensitive and specific diagnostics for the presence of toxigenic *C difficile* versus true disease is still in question

and will be the topic of this thesis. Much of the time, CDI cannot be clinically distinguished from other causes of diarrhea making laboratory diagnostics crucial for accurate diagnosis and treatment [13]. There is much discussion around the optimal way to conduct laboratory diagnosis of CDI regarding testing methods, sample selection, testing algorithms, and clinical validation [9, 10]. There exists no agreed upon gold standard (GS) for diagnostic testing of clinical CDI disease, and despite advances, no test combines high sensitivity/specificity, rapid turnaround time, and low cost [4, 5]. Current CDI diagnostics is based on a combination of both clinical and laboratory findings as current diagnostics are not able to distinguish between colonisation by the bacillus and true clinical CDI. Factors that are used for diagnosis include symptoms of diarrhea (3 or more unformed stools in a 24-hour period), positive laboratory diagnostic for presence of toxigenic *C. difficile* or its related toxins, or histopathologic findings of pseudomembranous colitis [5]. Due to high rates of colonization, it is crucial to only perform diagnostic tests on symptomatic patients after other causes of diarrhea such as laxatives, medications, and enteral feedings have been appropriately ruled out [5, 9]. Repeat testing increases the likelihood of a false positive results and diagnostics should not be used as a "test of cure" as results can remain positive for many weeks following culmination of the disease [5, 9].

There are multiple clinical diagnostic tests commercially available each with their own unique advantages, draw backs, and associated levels of detection. The following will be discussed: cell culture cytotoxicity neutralization assay (CCCNA), toxigenic culture, toxin enzyme immunoassays (EIA), glutamate dehydrogenase (GDH), and nucleic acid amplification tests (NAAT) (e.g., polymerase chain reaction (PCR) based tests).

Cell culture cytotoxicity neutralization assay (CCCNA) was historically considered to be the gold standard for diagnosis. It can detect down to picograms of *C. difficile* toxin but requires a long turnaround time (48 hours to 7 days), requires cell cultures (MCR-5, Vero, HeLa, Hep-2), requires trained personnel, and lacks uniform standardization [4, 9]. It functions by observing the cytopathic effects after a culture is inoculated with stool filtrates followed by specificity confirmation using a toxin B antitoxin for neutralization [4].

Toxigenic culture is performed by isolating *C. difficile* from the stool and confirming in culture if the isolate produces toxin [8, 9]. The turnaround time of >48 hours is not practical for routine clinical diagnostics and thus is not generally used in a clinical setting [9]. In 2010 the Infectious Disease Society of America (IDSA) supported this to be used for the gold standard in method comparison studies. Toxigenic culture will pick up more positive samples than CCCNA, however it has not been found superior to CCCNA in detection of clinical CDI disease [9].

Toxin enzyme immunoassays (EIAs) work by detecting toxin A, B, or both. These tests are inexpensive and yield rapid results with a high specificity (>95%) but have low sensitivity. Sensitivity using CCCNA as GS is 29-86% or using toxigenic culture as GS is 45-66% [9]. These tests are not recommended as a sufficient stand-alone test for CDI [4, 5, 9]. Growing evidence suggests the detection of free toxin in stool correlates best with clinical symptoms and outcomes [4]. Generally, it is considered that there is no stand-alone test that can accurately and quickly detect fecal free toxin for CDI diagnosis [14].

Glutamate dehydrogenase (GDH) is a metabolic enzyme produced by both toxigenic and nontoxigenic strains of *C. difficile* [4, 9]. GDH tests have high sensitivity, generally above 90% across most GS [9]. This is a rapid and inexpensive test that is commonly used for initial

screening (high negative predictive value) but will need to be followed by a second type of diagnostic, such as an EIA or NAAT detecting presence of toxin or a toxin gene to determine if a toxigenic strain is present [4, 5]. Combined assays that include GDH and EIA in a single lateral flow immunochromatographic membrane exist and utilize multiple tests in one platform to diagnose with variable results [4]. An example of this used in Alberta Public Laboratories is the TechLab C. Diff Quik Chek Complete that simultaneously tests for Toxin A/B and GDH.

Nucleic acid amplification tests (NAAT) became commercially available in 2009 and detect genes specific to toxigenic strains of *C. difficile* [4]. Popular NAAT options such as polymerase chain reaction (PCR) based tests have high sensitivity (85% to >95% depending on GS) and high specificity (>95%) for the presence of the organism or targeted gene [8, 9]. Targets for these PCR tests include conserved regions of tcdA (Toxin A Gene), tcdB (Toxin B Gene), and tcdC deletion (Binary toxin gene marker for identification of ribotype 027/NAP1 strain) [9, 15]. How to best utilize NAAT for the presence of toxigenic *C difficile* in diagnosing true disease is still in question and will be explored in the upcoming chapters. These NAATs pick up the presence of a gene that may or may not be expressed and their high level of detection cannot differentiate between asymptomatic carriage and true clinical disease. In 2014, 44% of US hospitals used NAAT as a stand-alone test for CDI diagnostics compared to Europe in 2013 at 5% [15]. In clinical settings where EIAs or CCCNA diagnostics are replaced by NAATs, a >50% increase in laboratory confirmed *C. difficile* can be observed. Patients colonized by *C. difficile* are 5 to 10 times more prevalent in hospital than patients with non-infectious diarrhea and/or CDI clinical disease [9, 13]. NAATs overcalling prevalence true disease can lead to misdiagnosis, unnecessary treatment, poor antibiotic stewardship, increased patient isolation, and unnecessary

clinical and economic burdens on healthcare facilities [15]. Further research is required to assess *C. difficile* NAATs for clinical disease detection and how to best implement this diagnostic tool.

Many facilities implement a multi-step algorithm for diagnostic testing that tries to address the limitations of the diagnostics and the complications presented by asymptomatic carriage of *C. difficile* [4, 5]. The European Society of Clinical Microbiology and Infectious Disease recommends a two-step algorithm utilizing an initially sensitive screening method (NAAT, or GDH), followed if positive by a more specific technique targeting toxin in stool (i.e., EIA, or NAAT) [4]. European guidelines recommend NAAT only be used as part of a two stage algorithm [15]. The intention of this method is to use an initial screening method with high negative predictive value followed by a diagnostic with sufficient positive predictive value in order to limit mis diagnosis as much as possible. The Infectious Disease Society of America has similar guidelines but also concludes that NAAT can be used alone if appropriate stool selection is guaranteed [4, 11].

1.3 Using PCR Cycle Threshold as a Predictor

Prior studies have tried to address PCR's inability to distinguish between true clinical disease and carriage of the *C. difficile* organism. Real-time PCR or quantitative PCR (qPCR) can be quantitative by an inverse correlation between the cycle threshold (Ct) with the log amount of the target sample [16]. Multiple studies have examined correlations between qPCR Ct value, free toxin and/or severity of CDI, or *C. difficile* bacterial fecal load [13, 17]. Identifying severity in many of these studies incorporate IDSA guidelines for leukocytes and creatinine with some

studies additionally factoring in temperature and evidence of severe colitis [13]. Study analyses commonly utilized receiver operating characteristic (ROC) curves and associated analyses attempting to define a single optimal cut-off value based on a standard chooses as a gold standard (GS) for the study. The following table briefly summarizes some of the attempts to use PCR Ct as a predictor for *C. difficile* severity and toxin load.

Authors	Findings
Garvey et al.	Ct value <26 correlated with ≥72% EIA positivity, a low Ct value is associated with higher mortality [13].
Senchyna et al.	Used PCR Ct value to predict free toxin. They found a similar cut-off of 26.35 predicted toxin positive samples with a sensitivity of 96.0% and specificity of 65.9% in addition to exploring other GS with similar results [14].
Crobach et al.	Found similar mean Ct values for Toxin negative symptomatic patients and asymptomatic carriers (30.4 [95% CI 29.5, 31.3] and 29.2 [95% CI 27.3, 31.2]) and a mean Ct of 24.4 [95% CI 23.5, 25.3] for toxin positive symptomatic patients. They recommend for clinical application establishing a Ct cut-off with a high negative predictive value of >90% (Ct= 29.0 or 32.0 in their data) could be useful to establish a preliminary diagnosis [18].
Kamboj et al.	Focused specifically on cancer patients to examine the potential of Ct to predict free toxin and clinically relevant disease. They found using CCCNA as GS an optimal Ct of 28 (using ROC and max Youden value) predicted a sensitivity of 77% and specificity of 74%. Lower Ct values were associated with toxin and increased disease severity [16].
Davies et al.	A retrospective study finding significantly different median Ct values for patents who died (Ct mean=25.5) versus survivors (Ct mean = 27.5) and between toxin positive (Ct mean= 24.9) and toxin negative (Ct mean = 31.6) patients. Optimizing their cut-off (ROC analysis) to Ct ≤25 with a sensitivity of 51.3% and specificity of 87.5%. for detection of toxin positive patients [15].
Reigadas et al.	Used Ct as a predictor of poor outcomes with the objective of developing a tool to better inform clinical plan and management of CDI [19]. Using ROC analysis, they found an optimal Ct of <23.5 predicted risk of poor outcomes in their combined cohort with a sensitivity of 67.4% and specificity of 87.3% [19].

Origuen et al.	Evaluated predictive merit of Ct with that of toxin detection and clinical variables for predicting severity of infection and outcomes by model comparison [20]. Their optimal Ct value of 27.55 yielded a sensitivity of 78.6% and specificity of 35.7% for poor outcomes. They found that including Ct into their predictive model did not meaningfully increase the predictive value over using toxin EIA and clinical variables alone. They observed a negative correlation with Ct value and disease severity but caution against clinical decisions on this value alone based on their modeling [20].
Sandlund et al.	Acknowledges the predictive ability of Ct values in the available literature but caution against solely using due to poor accuracy in toxin prediction and a significant overlap of Ct values between diseased and non-diseased patients. They argue for the use of ultrasensitive toxin assays instead of predictive Ct values as there appears to be a larger risk of misclassifying patients using current literature values of optimal Ct cut-off [12].

Table 1.1: Summary of past literature using PCR Ct as predictor for *C. difficile* toxin and disease.

The literature in Table 1.1 indicates it is possible to correlate *C. difficile* free toxin, and clinical outcomes with genomic burden with PCR cycle threshold. The literature has major shortcomings by attempting to simplify the disease and biologic process by seeking a single optimal Ct cut-point with PCR diagnostics to define disease and/or toxin load. These studies also have limited evaluative potential because a gold standard (GS) for CDI diagnostics does not exist, and few studies compare broadly Ct values across multiple GS and platforms. The literature however does set a foundation for further questions by hypothesizing PCR cycle Ct has potential to predict true clinical *C. difficile* infections (CDI) and that the high level of detection could be titrated to better represent presence of true clinical disease. The techniques used thus far have not provide a clinically relevant model. The publications fall short of providing a method that could be applied in a clinical setting as concluded by Sundlund et al., and Origuen et al. [12, 20]. ROC analyses and arithmetic means have been employed in many papers as a means of comparison for diagnostic tests but does not appear to capture the whole story of what we may be able to

extrapolate using genomic burden as an indicator. ROC curves have the benefit of allowing easy dichotomous test comparison but come with trade-offs. This method's common evaluation of the Area under the Curve (AUC), or area under the ROC curve, is simplistic but struggles with clinical application because it evaluates a test across all thresholds simultaneously; changes in sensitivity and specificity effect patients and diagnosis differently and in many cases might not be equally important [25]. Additionally, optimizing Ct values with a method such as Youden values, is a very basic statistical tool founded in mathematics and does not provide a clinical application to predict probability of disease.

This thesis builds upon the foundations of previous literature and conducts the first application of latent class analysis (LCA) to *C. difficile* diagnostics with the goal of using Ct to predict CDI to improve patient care. The use of LCA enables evaluation of diagnostics without the limitations of a single gold standard and allows one to evaluate probability of CDI more broadly in patients.

1.4 Summary of Subsequent Chapters

Chapter 2 establishes a method of CDI disease classification by PCR Ct value zones using a retrospective cohort of GDH positive patients and multiple diagnostic standards for comparison. Latent class analysis is utilized to better distinguish between symptomatic carriage of *C. difficile,* true CDI, and patients that have indeterminant diagnostic results. Latent class analysis allows for the use of more than one gold standard for comparison and eliminates reliance on a single gold standard for diagnostic evaluation. Through this process a classification system determining probability of CDI based on Ct values is established.

Chapter 3 examines a prospective cohort of patients with suspected CDI. The classification system established in the previous chapter based on Ct value zones is further evaluated for correlations with clinical metrics of disease severity.

Chapter 4 is the conclusion of the thesis discussion of its findings. This chapter expands on the possible clinical impacts, the value of what has been found, and presents future questions.

1.5 References

1. Katz KC, Golding GR, Choi KB, et al. The evolving epidemiology of *Clostridium difficile* infection in Canadian hospitals during a post epidemic period (2009-2015). Canadian Medical Association Journal 2018; 190(25): E758-E65.

2. Bouza E. Consequences of *Clostridium difficile* infection: understanding the healthcare burden. Clinical Microbiology and Infection 2012; 18: 5-12.

3. Mylonakis E, Ryan ET, Calderwood SB. *Clostridium difficile*-associated diarrhea – A review. Archives of Internal Medicine 2001; 161(4): 525-33.

4. Guery B, Galperine T, Barbut F. *Clostridioides difficile*: diagnosis and treatments. Bmj-British Medical Journal 2019; 366: 19.

5. Daniels LM, Kufel WD. Clinical review of *Clostridium difficile* infection: an update on treatment and prevention. Expert Opinion on Pharmacotherapy 2018; 19(16): 1759-69.

6. Levy AR, Szabo SM, Lozano-Ortega G, et al. Incidence and Costs of *Clostridium difficile* Infections in Canada. Open Forum Infectious Diseases 2015; 2(3): 10.

7. Bouza E, Munoz P, Alonso R. Clinical manifestations, treatment and control of infections caused by *Clostridium difficile*. Clinical Microbiology and Infection 2005; 11: 57-64.

8. Postma N, Kiers D, Pickkers P. The challenge of *Clostridium difficile* infection: Overview of clinical manifestations, diagnostic tools and therapeutic options. International Journal of Antimicrobial Agents 2015; 46: S47-S50.

9. Korman TM. Diagnosis and Management of *Clostridium difficile* Infection. Seminars in Respiratory and Critical Care Medicine 2015; 36(1): 31-43.

10. Loo VaDIa. Association of Medical Microbiology and Infectious Disease Canada treatment practice guidelines for *Clostridium difficile* infection. Official Journal of the Association of Medical Microbiology and Infectious Disease Canada 2018; 3: 71-92.

11. McDonald LC, Gerding DN, Johnson S, et al. Clinical Practice Guidelines for *Clostridium difficile* Infection in Adults and Children: 2017 Update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). Clinical Infectious Diseases 2018; 66(7): 987-94.

12. Sandlund J, Wilcox MH. Ultrasensitive Detection of Clostridium difficile Toxins Reveals Suboptimal Accuracy of Toxin Gene Cycle Thresholds for Toxin Predictions. Journal of Clinical Microbiology 2019; 57(6): 5.

13. Garvey MI, Bradley CW, Wilkinson MAC, Holden E. Can a toxin gene NAAT be used to predict toxin EIA and the severity of *Clostridium difficile* infection? Antimicrobial Resistance and Infection Control 2017; 6: 8.

14. Senchyna F, Gaur RL, Gombar S, Truong CY, Schroeder LF, Banaei N. *Clostridium difficile* PCR Cycle Threshold Predicts Free Toxin. Journal of Clinical Microbiology 2017; 55(9): 2651-60.

15. Davies KA, Planche T, Wilcox MH. The predictive value of quantitative nucleic acid amplification detection of *Clostridium difficile* toxin gene for faecal sample toxin status and patient outcome. Plos One 2018; 13(12): 9.

16. Kamboj M, Brite J, McMillen T, et al. Potential of real-time PCR threshold cycle (C-T) to predict presence of free toxin and clinically relevant *C. difficile* infection (CDI) in patients with cancer. Journal of Infection 2018; 76(4): 369-75.

17. Dionne LL, Raymond F, Corbeil J, Longtin J, Gervais P, Longtin Y. Correlation between *Clostridium difficile* Bacterial Load, Commercial Real-Time PCR Cycle Thresholds, and Results of Diagnostic Tests Based on Enzyme Immunoassay and Cell Culture Cytotoxicity Assay. Journal of Clinical Microbiology 2013; 51(11): 3624-30.

18. Crobach MJT, Duszenko N, Terveer EM, Verduin CM, Kuijper EJ. Nucleic Acid Amplification Test Quantitation as Predictor of Toxin Presence in *Clostridium difficile* Infection. Journal of Clinical Microbiology 2018; 56(3).

19. Reigadas E, Alcala L, Valerio M, Marin M, Martin A, Bouza E. Toxin B PCR cycle threshold as a predictor of poor outcome of *Clostridium difficile* infection: a derivation and validation cohort study. Journal of Antimicrobial Chemotherapy 2016; 71(5): 1380-5.

20. Origuen J, Orellana MA, Fernandez-Ruiz M, et al. Toxin B PCR Amplification Cycle Threshold Adds Little to Clinical Variables for Predicting Outcomes in *Clostridium difficile* Infection: A Retrospective Cohort Study. Journal of Clinical Microbiology 2019; 57(2).

21. Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: the case of tests with continuous results. Biochemia Medica 2016; 26(3): 297-307.

22. Habibzadeh F, Habibzadeh P. The likelihood ratio and its graphical representation. Biochemia Medica 2019; 29(2): 7.

23. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KGM, de Groot JAH. Latent Class Models in Diagnostic Studies When There is No Reference Standard-A Systematic Review. American Journal of Epidemiology 2014; 179(4): 423-31.

24. Visser I, Speekenbrink M. depmixS4: An R Package for Hidden Markov Models. Journal of Statistical Software 2010; 36(7): 1-21.

25. Halligan S, Altman DG, Mallett S. Disadvantages of using the area under the receiver operating characteristic curve to assess imaging tests: a discussion and proposal for an alternative approach. Eur Radiol 2015; 25:932–9.

2.0 Latent Class Analyses for the Diagnosis of *Clostridioides difficile* Infection

Published in Clinical infectious diseases: an official publication of the Infectious Diseases Society of America 2020.

Authors: Cody P. Doolan, Thomas Louie, Christopher Lata, Oscar E. Larios, William Stokes, Joseph Kim, Kristen Brown, Paul Beck, Rob Deardon, and Dylan R. Pillai

2.1 Chapter Abstract

Background. *Clostridioides difficile* infection (CDI) is an opportunistic disease that lacks a gold-standard test. Nucleic acid amplification tests such as real-time polymerase chain reaction (PCR) demonstrate an excellent limit of detection (LOD), whereas antigenic methods are able to detect protein toxin. Latent class analysis (LCA) provides an unbiased statistical approach to resolving true disease.

Methods. A cross-sectional study was conducted in patients with suspected CDI (N = 96). Four commercial real-time PCR tests, toxin antigen detection by enzyme immunoassay (EIA), toxigenic culture, and fecal calprotectin were performed. CDI clinical diagnosis was determined by consensus majority of 3 experts. LCA was performed using laboratory and clinical variables independent of any gold standard.

Results. Six LCA models were generated to determine CDI probability using 4 variables including toxin EIA, toxigenic culture, clinical diagnosis, and fecal calprotectin levels. Three defined zones as a function of real-time PCR cycle threshold (Ct) were identified using LCA: CDI likely (>90% probability), CDI equivocal (<90% and >10%), CDI unlikely (<10%). A

single model comprising toxigenic culture, clinical diagnosis, and toxin EIA showed the best fitness. The following Ct cutoffs for 4 commercial test platforms were obtained using this model to delineate 3 CDI probability zones: GeneXpert®: 24.00, 33.61; Simplexa®: 28.97, 36.85; Elite MGB®: 30.18, 37.43; and BD Max™: 27.60, 34.26.

Conclusions. The clinical implication of applying LCA to CDI is to report Ct values assigned to probability zones based on the commercial real-time PCR platform. A broad range of equivocation suggests clinical judgment is essential to the confirmation of CDI.

2.2 Introduction

Clostridioides difficile is the predominant cause of nosocomial and healthcare-associated infectious diarrhea in the community [1, 2]. *Clostridioides difficile* manifestation ranges from asymptomatic colonization to severe fulminant disease and/or death [3]. Studies have shown up to 15% of healthy adults and up to 21% of hospital-admitted patients can be asymptomatically colonized with *C. difficile* [4]. The transition from colonization to clinical disease is multifactorial, dependent on types and durations of antimicrobial exposure, agents that alter intestinal physiology, and host factors including intestinal immunity and bile salt metabolism. Asymptomatic colonization and transition to clinically evident infection can be bidirectional where patients with CDI can be self-resolving or run a relapsing course. Because of these complex pathogen, host, and microbial dynamics, detecting the presence of *C. difficile* toxin in stool by either antigenic or molecular methods does not directly correlate with clinical disease [5].

Accurately diagnosing CDI with a stand-alone test remains elusive and available tests come with trade-offs [6–8]. The presence of *C. difficile* toxin correlates well with clinical disease but enzyme immunoassays (EIAs), although simple and inexpensive, lack sufficient analytical sensitivity to be used alone [6, 8]. Nucleic acid amplification tests (NAATs), using real-time polymerase chain reaction (PCR) targeting a toxin-producing gene, provide high analytical sensitivity for the detection of *C. difficile* bacteria at the cost of specificity for clinical disease, especially without proper assessment of pretest probability and indiscriminate test utilization to help distinguish between clinical CDI and asymptomatic carriage and colonization. Despite its excellent limit of detection (LOD), PCR has not traditionally been used to distinguish between clinical CDI and asymptomatic carriage and colonization [2–4, 6, 8]. Detection of glutamate dehydrogenase (GDH), a metabolic enzyme produced by both toxigenic and nontoxigenic strains, provides inexpensive and rapid screening but must be followed with a confirmatory EIA or NAAT [8]. Fecal calprotectin (fCPT) has also been demonstrated to correlate with CDI and the presence of toxigenic *C. difficile* [9, 10]. Fecal calprotectin is a stable protein that is present in proportion to neutrophil migration in the mucosa of the intestine, thus acting as a marker for CDI- and non– CDI-related inflammatory bowel disease [9, 10].

Generally, a 2-step algorithm is recommended for CDI diagnosis, but some guidelines do state that NAATs are sufficient as a stand-alone test [8, 11] supported with clinical context. Several hospitals have reported a 50–100% increase in the amount of laboratory-confirmed CDI when switching from EIA to stand-alone NAATs [12]. We posed the question: Can the exceptional LOD of NAATs through the use of cycle threshold (Ct) values be used as a means to better identify clinically significant CDI in situations where noninfectious diarrhea can occur?

Several groups have attempted to use the quantitative potential of real-time PCR to address this concern. It has been demonstrated that reporting the real-time PCR Ct can reduce the treatment of toxin-negative patients [13]. Ct values have an inverse correlation with the log amount of the target gene and thus bacterial load [14, 15]. Laboratories have demonstrated that Ct values strongly correlate with fecal protein toxin [11, 14, 16–19]. Ct values have been correlated with CDI clinical outcomes and disease severity [14, 17, 18, 20, 21]. Most of the literature optimizing CDI real-time PCR Ct values has sought a single cutoff to determine disease status through the use of receiver operator curves (ROCs) [16, 17]. ROC analysis can be used to assess a discriminating threshold and test performance but is limited to a single reference method and single optimal cutoff [22]. So far, these efforts have failed to yield a method that has widespread clinical validity for predicting clinical CDI and/or protein toxin load in order to reduce laboratory detection of colonization. Commercial assays typically use microbiological endpoints to determine Ct value cutoffs [23, 24]. We sought to use latent class analysis (LCA) as an unbiased approach to determine CDI status based on multiple molecular real-time PCR platforms, toxin antigen detection by EIA, the clinical judgment of experts, and markers of inflammation namely fCPT.

2.3 Methods

Study Design, Patient Population, and Ethics

A subset of patients (N = 96) with suspected CDI who tested positive for the presence of GDH were selected at random during the month of October 2016 to November 2016 in Calgary, Alberta. All samples were received at Alberta Precision Laboratories, the primary microbiology laboratory for the city of Calgary covering a population of approximately 1.4 million, comprising both inpatients and outpatients. Alberta Precision Laboratories' CDI clinical diagnostic algorithm relies on a 2-step algorithm. Stool samples were first tested by Liaison *C. difficile* GDH (DiaSorin, Cypress, CA) for GDH. Samples positive for GDH were then tested with GeneXpert *C. difficile* (Cepheid, Sunnyvale, CA) real-time PCR to report tcdB presence. Ethical approval was obtained through University of Calgary's Institutional Research Information Services Solution (REB-16–1896). Patient information was protected as per Alberta Health Services protocol and all data were de-identified prior to analysis.

Additional Diagnostic Testing

Additional diagnostic testing and clinical chart review beyond the routine were conducted on the samples from the 96 patients testing positive for GDH (Figure 2.1). Additional diagnostics included 3 additional commercial NAAT real-time PCR platforms: Simplexa C. difficile Direct (DiaSorin, Cypress, CA), *C. difficile* Elite MGB kit (ELITechGroup S.p.A, Torino, Italy), and BD Max *C. difficile* assay (Becton Dickinson, Sparks, MD). Liaison *C. difficile* Toxins A&B (DiaSorin, Stillwater, MN) was used to evaluate the presence of toxin A/B (Toxin). Toxigenic culture using standard anaerobic culture on selective cycloserine cefoxitin fructose agar (CCFA; Dalynn Biological, Alberta, Canada) was performed to further evaluate the presence of the organism. *Clostridioides difficile* isolates were confirmed by matrix-assisted laser desorption

ionization–time of flight mass spectrometry (MALDI-TOF; Vitek MS; Biomerieux, Durham, NC) and shown to be toxigenic strains by performing real-time PCR to determine the presence of the tcdB gene using Cepheid GeneXpert *C. difficile*. fCPT was measured using Liaison calprotectin (DiaSorin, Stillwater, MN), a quantitative sandwich EIA relying on a monoclonal antibody against the fCPT heterocomplex. fCPT is a nonspecific marker of intestinal inflammation associated with CDI [9, 10].

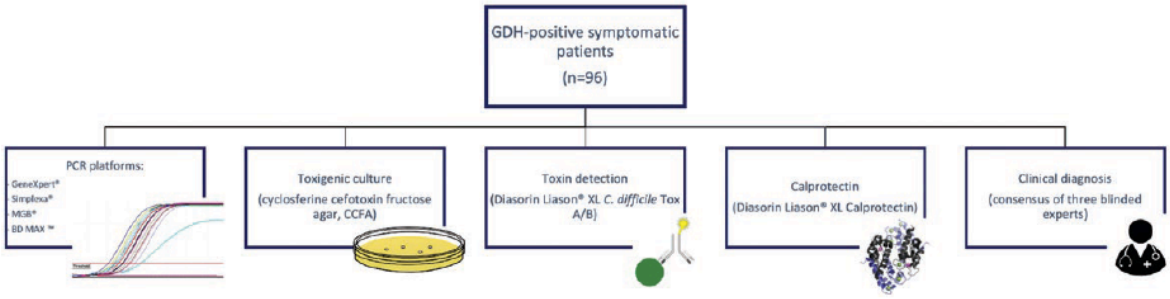


Figure 2.1: Flow chart demonstrating criteria for patient enrollment and subsequent laboratory and clinical investigations performed. Abbreviations: *C. difficile*, *Clostridioides difficile*; GDH, glutamate dehydrogenase; PCR, polymerase chain reaction. Images are licensed under the Creative Commons Attribution 4.0 International license.

Clinical Diagnostic Evaluation by Experts

Clinical diagnosis of CDI was determined by conducting a chart review and itemizing other relevant diagnostic results. For a list of clinical and laboratory data gathered from the chart review to form the diagnostic verdict, see Appendix Table 2.1. A positive clinical diagnosis was based on the consensus majority of 3 independent, blinded, infectious disease experts. Latent class analysis for diagnosing disease that lacks a single gold-standard laboratory test, multiple diagnostic tests can be combined into a latent class model to determine the probability of true disease. Latent class analysis permits probabilistic determination of an unmeasured disease "latent class" using model inputs of multiple "observed variables" [25, 26]. Multiple observed variables were investigated in the following combinations: A, clinical diagnosis and toxin; B,

clinical diagnosis and culture; C, clinical diagnosis, toxin, and culture; D, clinical diagnosis and fCPT; E, clinical diagnosis, toxin, and fCPT; F, clinical diagnosis, toxin, culture, and fCPT (Figure 2.2). Clinical diagnosis was used in each model as CDI in practice is ultimately determined by healthcare provider judgment [7]. Bayesian information criterion (BIC) values, a measure of fitness of the LCA model, were calculated for each model to aid in model comparison. A lower BIC implies better LCA model fitness [27]. Latent class analyses were conducted in R-Studio (version 1.2.5033; R Foundation for Statistical Computing, Vienna, Austria) using the package depmixS4 [28, 29]. The mean probability of membership to a disease latent class as a function of Ct value was plotted and analyzed via polynomial regression for each commercial real-time PCR assay. The corresponding adjusted R^2 values were evaluated to measure the fit of each regression model. Models with an adjusted R^2 value greater than 0.5 were deemed sufficiently predictive. The model that incorporated the most clinically relevant variables, has an adjusted R^2 greater than 0.5, and lower BIC was chosen to ascertain Ct cutoff values for disease probability. Cutoff values for disease and non-disease status were set at 90% and 10% probability of disease in a Ct range of 0 to 40 in line with the commercial real-time PCR manufacturer's recommendations. The LCA models had Ct values divided into 3 groups: CDI likely (>90% probability of disease), CDI equivocal (10% and 90% probability of disease), CDI unlikely (<10% probability of disease). The LCA described was performed for each real-time PCR platform.

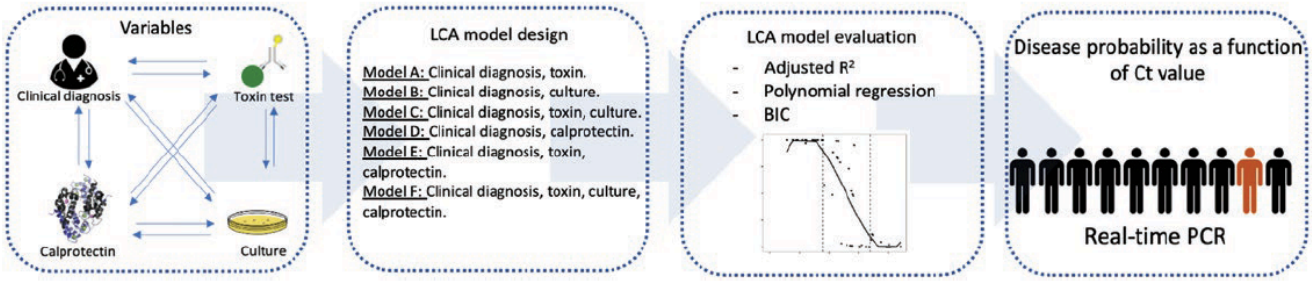


Figure 2.2: Overview of key variables used to generate latent class analysis model, statistical tools used to select model fitness, and outputs generated to determine likelihood of *Clostridioides difficile* infection. Abbreviations: BIC, Bayesian information Criterion; Ct. Cycle Threshold; LCA, Latent class analysis; PCR, polymerase chain reaction. Images are licensed under the Creative Commons Attribution 4.0 International license.

2.4 Results

Patient Demographics

In total, 96 patients with diarrhea suspected of CDI and GDH positive by EIA were included in the study. Among this cohort of patients, 48.9% were female, 52.0% were treated for CDI, and 60.4% of the patients had a record of antibiotic use in the prior 3 months. The average patient age was 48.5 years (1–91 years) with a median age of 50.5 years (Table 2.1).

Real-time PCR Assays, Toxigenic Culture, Clinical Diagnosis, and Fecal Calprotectin

The distribution of Ct values for each patient sample in the study across all commercial molecular platforms is shown in Figure 2.3. The positivity rates between commercial real-time PCR platforms (N = 96) using the manufacturer's recommended Ct value cutoff were as follows: GeneXpert, 68.4% positive; Simplexa, 66.3% positive; BD Max, 61.1% positive; and Elite MGB, 66.3% positive (Table 2.1). Toxigenic culture using CCFA was positive in 74.7% of patients. The consensus clinical diagnosis of CDI by 3 infectious disease specialists based on selected criteria (Appendix Table 2.1) was made in 52.6% of patients. The toxin antigens A/B

were detected in 41.1% of patients using EIA. The mean concentration of the inflammatory marker fCPT was 316.9 µg/g with a range from 5 to 800 µg/g (Table 2.1).

Table 2.1: Patient Characteristics, Risk Factors for *Clostridioides difficile* Infection, and Results of Clinical Laboratory Investigation in the Study Population

Variables	Percentage	Mean	Range
Age, y	…	48.46	1 to 91
Females	48.9	…	…
Previous CDI	21.6	…	…
Antibiotics in previous 6 m	65.3	…	…
Antibiotics in previous 3 m	60.4	…	…
Antibiotics in previous week	32.7	…	…
Treated for CDI	52	…	…
Clinical diagnosis (+)	52.6	…	…
Culture (+)	74.7	…	…
Toxin detected by EIA (+)	41.1	…	…
fCPT, µg/g	…	316.9	5 to 800
GeneXpert® (+)	68.4	…	…
Simplexa® (+)	66.3	…	…
BD Max™ (+)	61.1	…	…
Elite MGB® (+)	66.3	…	…

N = 96.
Abbreviations: CDI, *Clostridioides difficile* infection; EIA, enzyme immunoassay; fCPT, fecal calprotectin; +, positive.

Latent Class Analyses to Determine *Clostridioides difficile* Infection

Latent class analysis was used to investigate the predictive ability of real-time PCR Ct values to identify CDI without being limited to a single gold-standard reference method. Probability of membership to 1 of 3 latent classes (CDI, not CDI, indeterminant) as a function of Ct value was evaluated for 6 different models of the observed variables (clinical diagnosis [clinical diagnosis], toxin detection by EIA [toxin], toxigenic culture [culture], and fCPT detection by EIA [calprotectin]), all as a function of Ct value for each commercial real-time PCR platform. The statistical values associated with LCA model fitness for all models (A through F) are summarized in Appendix Table 2.2. All models are presented in the Appendix Figures 2.1–2.4. Model A (toxin, clinical judgment) resulted in a high adjusted R^2 and relatively low BIC. Model B (culture, clinical judgment) did not consistently meet the criteria of an adjusted R^2 greater than

0.5 but resulted in a lower BIC value relative to the other models generated. With the introduction of fCPT, in models D, E, and F, BIC values were consistently much higher (Appendix Table 2.2). Model C (when both toxin and culture are combined with clinical judgment) incorporated the greatest diversity of observed variables while still maintaining an adjusted R^2 greater than 0.5 and comparatively lower BIC values. Latent class analysis results for model C showing the probability of CDI in relation to the Ct values for 4 commercial real-time PCR systems used in this study are presented in Figure 2.4. The LCA model C generated the following Ct cutoff values: GeneXpert, an average 90% probability of disease Ct value cutoff was 24.00 and an average 10% probability of disease Ct value cutoff was 33.61; Simplexa, an average 90% probability of disease Ct value was 28.97 and an average 10% probability of disease Ct value 36.85; Elite MGB, an average 90% probability of disease Ct value was 30.18 and an average 10% probability of disease Ct value 37.43; and BD Max, an average 90% probability of disease Ct value was 27.60 and an average 10% probability of disease Ct value 34.26. All models generated comparable Ct value cutoffs except for models B and D (Appendix Table 2.2).

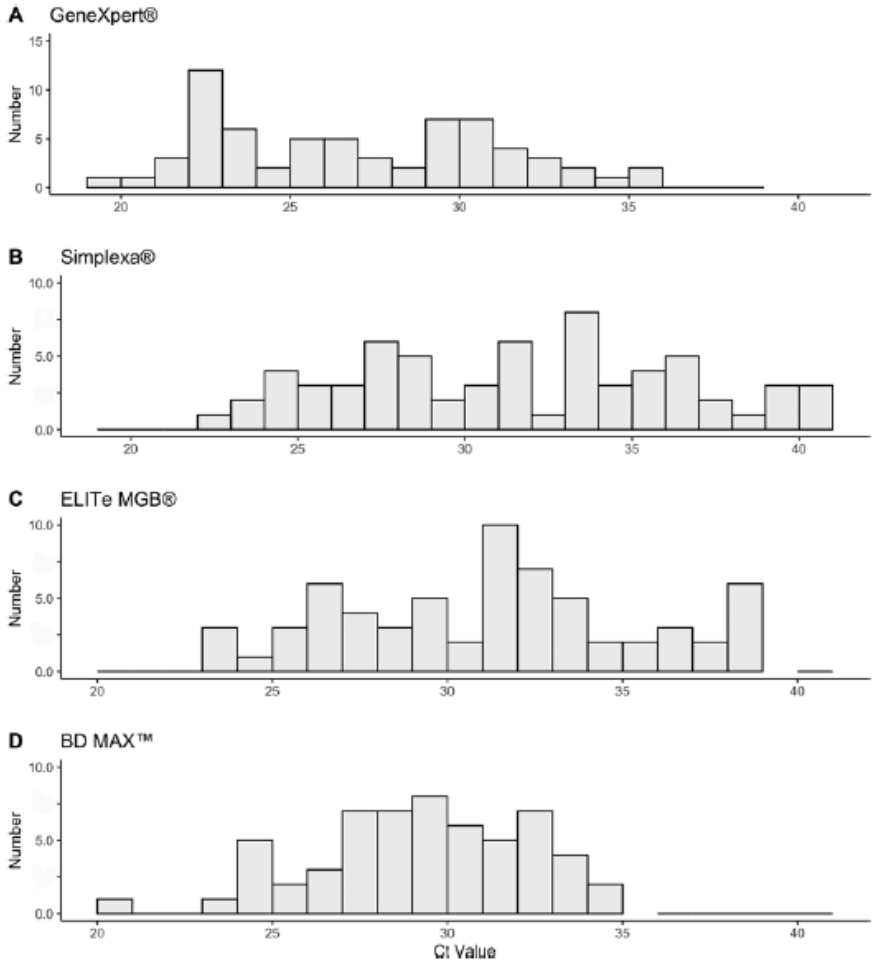


Figure 2.3 Histogram plot of Ct values for each commercial real-time PCR platform designed to detect the *Clostridioides difficile* toxin gene. Ct values are plotted against number for each of (A) Cepheid GeneXpert®, (B) DiaSorin Simplexa®, (C) Becton Dickinson BD Max™, and (D) ELITechGroup Elite MGB®. Abbreviations: Ct, cycle threshold; PCR, polymerase chain reaction.

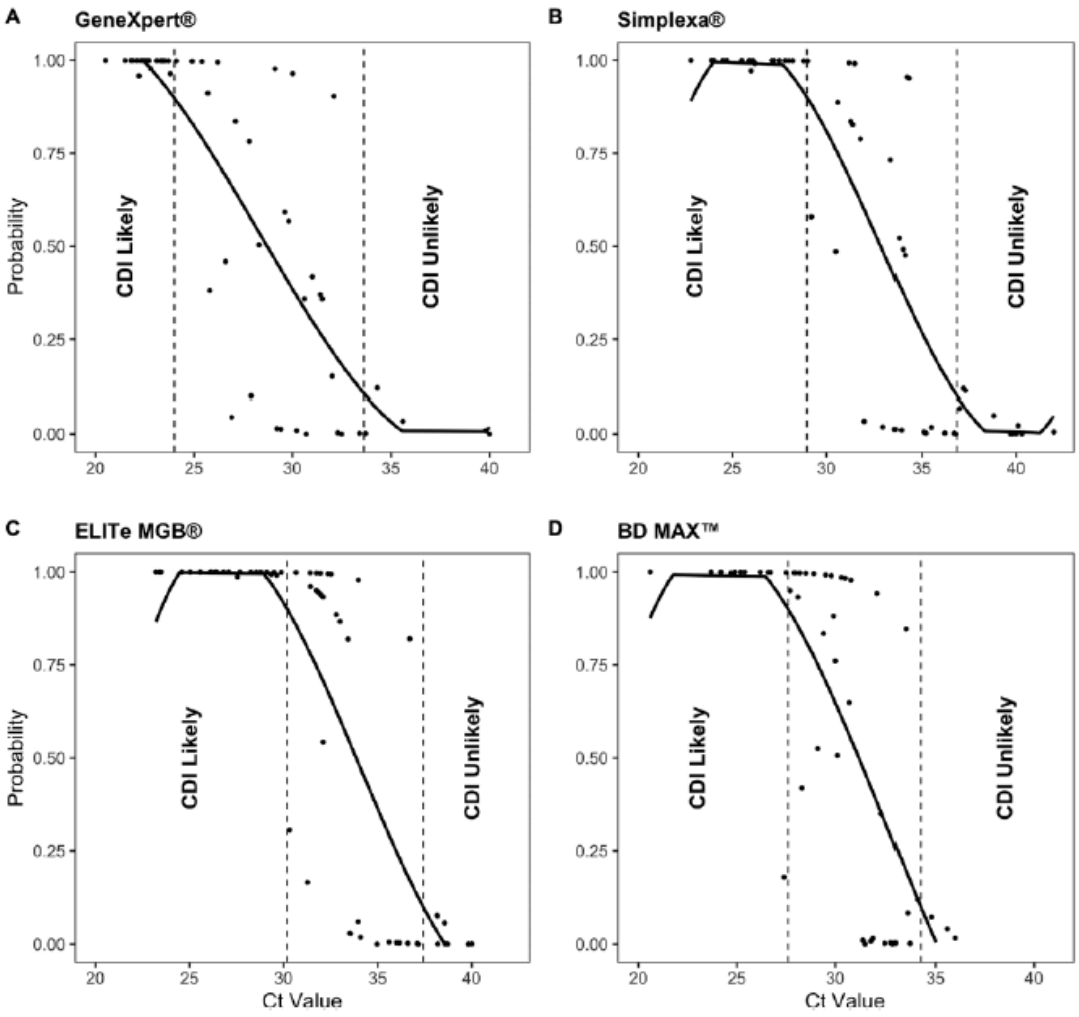


Figure 2.4. Latent class analyses results of model C showing the probability of CDI in relation to the Ct value (N = 96) for commercial real-time PCR systems used in this study: Cepheid GeneXpert®, DiaSorin Simplexa® real-time PCR, ELITe MGB® real-time PCR, and BD Max™ real-time PCR. Observed variables included in model C: clinical diagnosis (diagnosis), toxin detection by enzyme immunoassay (toxin), toxigenic culture (culture). Vertical dashed lines demarcate the Ct value boundaries obtained for the equivocal zone based on 10% and 90% probability of CDI using polynomial regression. Other models are presented in the Appendix Data (Table 2, and Figures 2.1-2.4). Abbreviations: CDI, *Clostridioides difficile* infection; Ct, cycle threshold; PCR, polymerase chain reaction.

2.5 Discussion

Given no gold-standard reference test for CDI diagnosis, LCA combines multiple standards into a single model. The model focuses on major components identified as necessary for clinical disease: the presence of *C. difficile* organism, presence of protein toxin, clinical symptoms as

adjudicated by experts, and a heightened immune response. These factors were then correlated with Ct values produced by commercial molecular platforms commonly used to test for CDI. The LCA model investigates combinations of observed variables in order to define the probability of CDI. The model where clinical diagnosis, protein toxin, and culture of the organism are combined produced the most robust model from a statistical perspective. fCPT in this model did not add additional predictive value and was excluded in the final model of choice. The higher BIC values (poor LCA fitness) in models incorporating fCPT are likely due to high variability in calprotectin values in patients who have CDI toxin gene detectable by real-time PCR [9, 10]. Taken together, the LCA suggests a relatively large zone of equivocal diagnosis based on Ct value for all commercial platforms evaluated in several of the models generated. Using a combination of statistical tools, a single model appeared to provide the most complete picture of CDI—namely, with toxigenic culture, clinical diagnosis by experts, and protein toxin included. This suggests that all 3 aforementioned variables add value to the diagnosis of CDI and real-time PCR Ct values demarcate an equivocal range based on the commercial real-time PCR assay where more information is required. Previous attempts to find an optimal Ct cutoff for toxin or CDI generally place their Ct values within the equivocal zones defined by the LCA performed in this study. The broad range of equivocal findings in the LCA likely explains why a single cutoff has poor predictive value for the clinical diagnosis of CDI. Although results for patients falling within the equivocal zone do not gain more resolution with the current LCA, it provides value in defining when an accurate stand-alone laboratory result for clinical disease is not possible. Other infectious disease diagnostic studies have benefited from LCA modeling [30–36]. For example, we recently reported that antigenic tests and molecular test results can differ from the clinical diagnosis of acute dengue infection, and that LCA is able to provide an

unbiased approach to determining true acute disease when the diagnostic tests do not agree [34]. There is immediate potential for clinical application of probability zones: CDI likely, CDI equivocal, and CDI unlikely. The probability models can provide diagnostic guidance that is more insightful than the current binary real-time PCR result. LCA applied to commercial NAAT diagnostic technology to provide clinicians with additional valuable information at little to no additional cost to hospitals or laboratories. In 2014, 44% of US hospitals used NAAT as a stand-alone test for CDI diagnostics [17]. The LCA model developed here can reduce the overdiagnosis of CDI, potentially leading to improved antibiotic stewardship and increased targeted treatment for patients. Analyses incorporated multiple standards into one model but was not powered to account for variations such as age or comorbidities of the patients; for example, an older immunocompromised patient might present with clinically relevant CDI at a significantly higher Ct value. Further studies integrating additional immune biomarkers might increase model resolution in the equivocal zones. There is an assumption of conditional independence with the observed variables of the LCA, but all observed variables address clinical symptoms and/ or pathophysiology of the same disease processes. Additionally, an assumption was made that more variables used in the LCA, an adjusted R^2 greater than 0.5, and lower BIC value implied greater model fitness to predict CDI. While it is difficult to say that one set of criteria produces the "correct" model, the use of an adjusted R^2 greater than 0.5 and lower BIC value is a conservative approach, lowering the chance of overfitting and including terms that only appear important due to random chance. Further investigation correlating clinical outcomes and/or disease severity to the Ct value cutoff zones proposed here will be helpful in further clinically validating the LCA model.

Conclusions

This publication applied an unbiased approach to CDI diagnosis relying on LCA and multiple standards to determine disease status. The results show that CDI diagnosis using a single Ct value cutoff likely oversimplifies diagnosis. A broad range of equivocation exists where clinical diagnosis is mandatory. Commercial platforms also do not necessarily have identical cutoff points for zones of CDI probability, suggesting heterogeneity between assays. The LCA approach has implications for infectious disease diagnostics beyond CDI where pathogen gene presence does not necessarily imply clinical disease.

2.6 Notes

Acknowledgments

The authors thank the expert technical assistance of medical laboratory technologists at Alberta Public Laboratories. We also thank Daniel Castaneda Mogollon for expertise with graphics generated for the figures. Financial support. This work was supported by Canadian Institutes for Health Research (PJT-156251) and Cumming School of Medicine (Clinical Research Fund), University of Calgary. Laboratory test kits were provided by the commercial manufacturers free of charge. Potential conflicts of interest. T. L. reports being an investigator for Summit PLC, Vedanta, Seres Therapeutics, Finch, Immunimed, Rebiotix, and MGB Biopharma and consulting for Davoterra, outside the submitted work. All other authors report no potential conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

2.7 Chapter 2 Appendix Figures and Tables

Appendix Table 2.1: Summary of variables available to infectious disease experts to make clinical diagnosis based on chart review. Abbreviations in table: IBD – diagnosis of inflammatory bowel disease. CRP – C reactive protein. WBC- white blood cell count. Hgb – hemoglobin. CT – computed tomography. CDI – *C. difficile* infection.

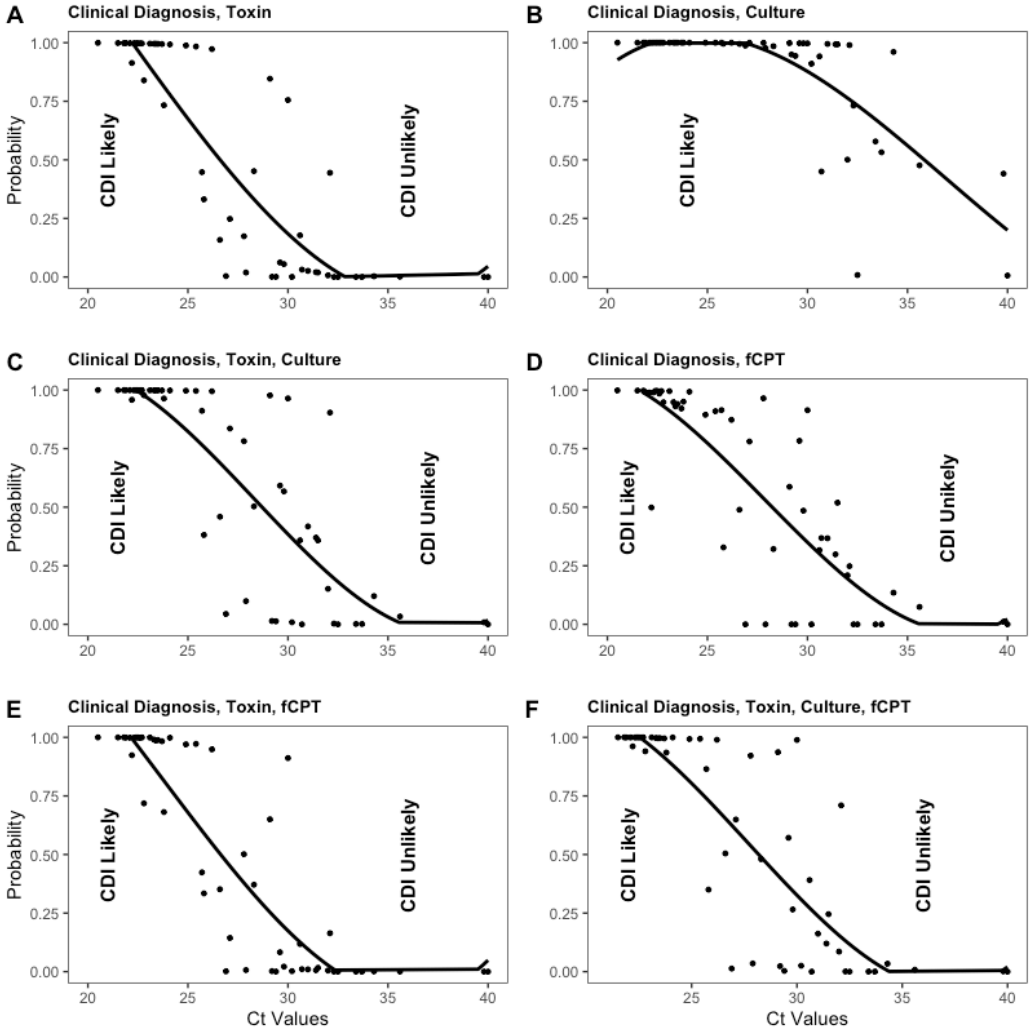
Patient Info	Clinical	Other	Laboratory
Test Date	Antibiotic Used	X-Ray	Stool WBC
Age	Treated for CDI	CT Scan	Hgb
Sex	Repeated Treatment	Endoscopy	WBC
In/Outpatient	Admitted	Biopsy	Platelets
Previous CDI	Fever (>38C)	IBD	CRP
Antibiotics in last 3 Months	Tachycardia		
Antibiotics in last 6 Months			
Antibiotics in last week			
Referred for fecal transplantation			
Admitting Diagnosis			

Appendix Table 2.2: Results from latent class analyses for all four commercial real-time PCR platforms evaluated in this study and associated statistical values corresponding to the six models (A through F) generated. Abbreviations: BIC – Bayesian Information Criterion. Ct – cycle threshold.

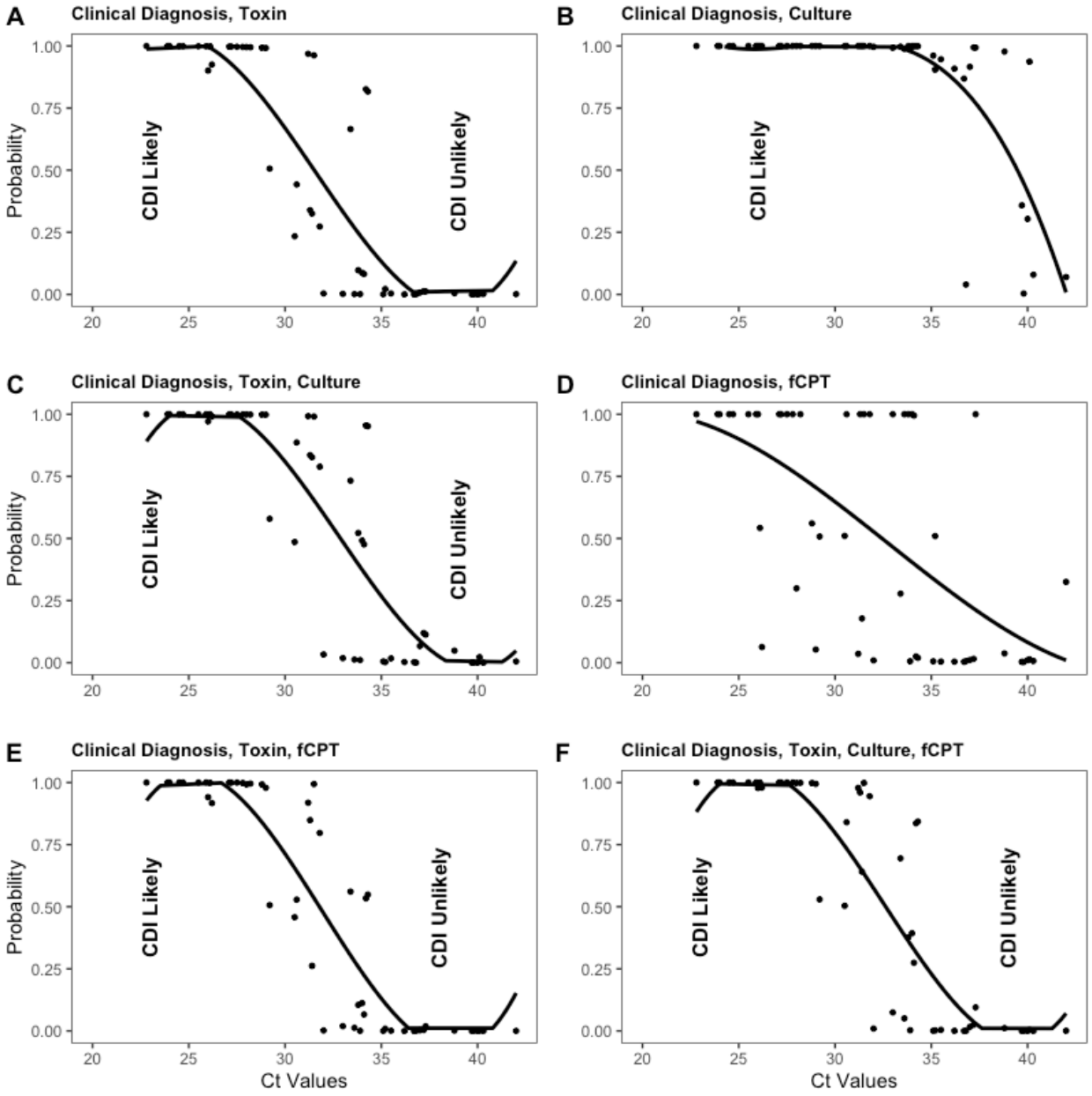
		GeneXpert®				Simplexa®			
Observed Variables	**LCA model**	**BIC**	**Adjusted R-Squared**	**Ct at 90% Probability of Disease**	**Ct at 10% Probability of Disease**	**BIC**	**Adjusted R-Squared**	**Ct at 90% Probability of Disease**	**Ct at 10% Probability of Disease**
Clinical Diagnosis, Toxin	A	207.2	0.761	22.98	31.12	207.4	0.761	27.58	35.38
Clinical Diagnosis, Culture	B	179.6	0.566	29.46	-	185.1	0.604	35.62	41.56
Clinical Diagnosis, Toxin, Culture	C	297.6	0.659	24.00	33.61	298.3	0.746	28.97	36.85
Clinical Diagnosis, fCPT	D	1492.7	0.673	23.26	33.39	1411.1	0.342	-	-
Clinical Diagnosis, Toxin, fCPT	E	1590.6	0.781	22.97	30.93	1586.5	0.825	28.14	35.32
Clinical Diagnosis, Toxin, Culture, fCPT	F	1681.2	0.702	23.89	32.63	1677.7	0.777	28.90	36.37

		ELITe MGB®				BD Max™			
Observed Variables	**LCA model**	**BIC**	**Adjusted R-Squared**	**Ct at 90% Probability of Disease**	**Ct at 10% Probability of Disease**	**BIC**	**Adjusted R-Squared**	**Ct at 90% Probability of Disease**	**Ct at 10% Probability of Disease**
Clinical Diagnosis, Toxin	A	200.3	0.712	29.56	36.96	214.8	0.549	26.82	33.74
Clinical Diagnosis, Culture	B	175.8	0.686	31.85	39.41	200.0	0.367	-	-
Clinical Diagnosis, Toxin, Culture	C	284.1	0.684	30.18	37.43	305.9	0.558	27.60	34.26
Clinical Diagnosis, fCPT	D	1403.3	0.424	-	-	1411.4	0.263	-	-
Clinical Diagnosis, Toxin, fCPT	E	1580.6	0.801	28.96	35.27	1598.2	0.587	26.81	33.57
Clinical Diagnosis, Toxin, Culture, fCPT	F	1664.4	0.796	30.64	36.95	1689.0	0.610	27.86	34.05

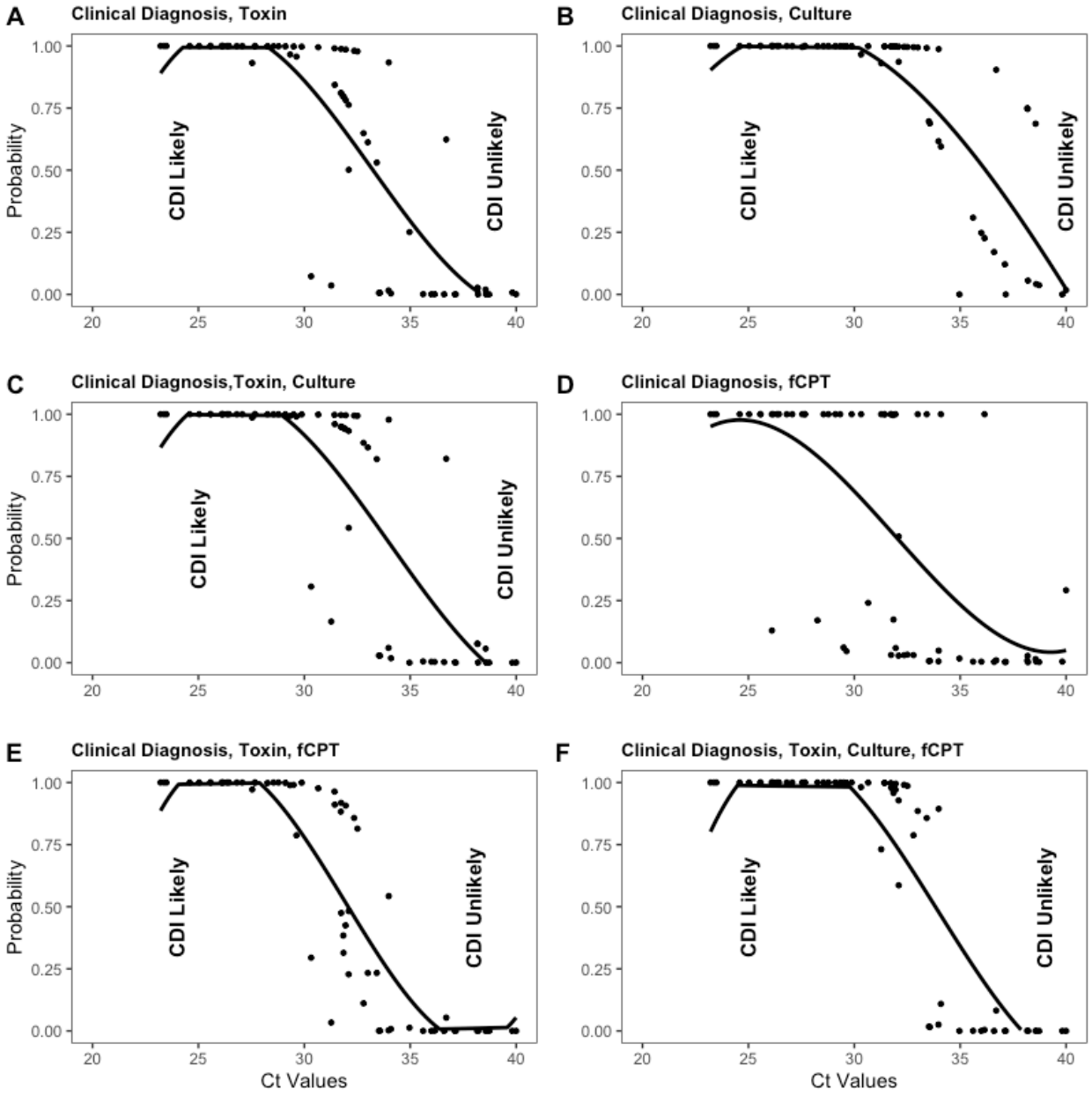
Appendix Figure 2.1: LCA analyses results depicting probability of CDI in relation to the Ct value for Cepheid GeneXpert® qPCR. Panels A-F represent each variable: clinical diagnosis, toxin detection by DiaSorin Liaison® XL *C. difficile* Toxin A/B [Toxin], toxigenic culture [Culture], and detection of calprotectin as a maker of inflammation [fCPT].



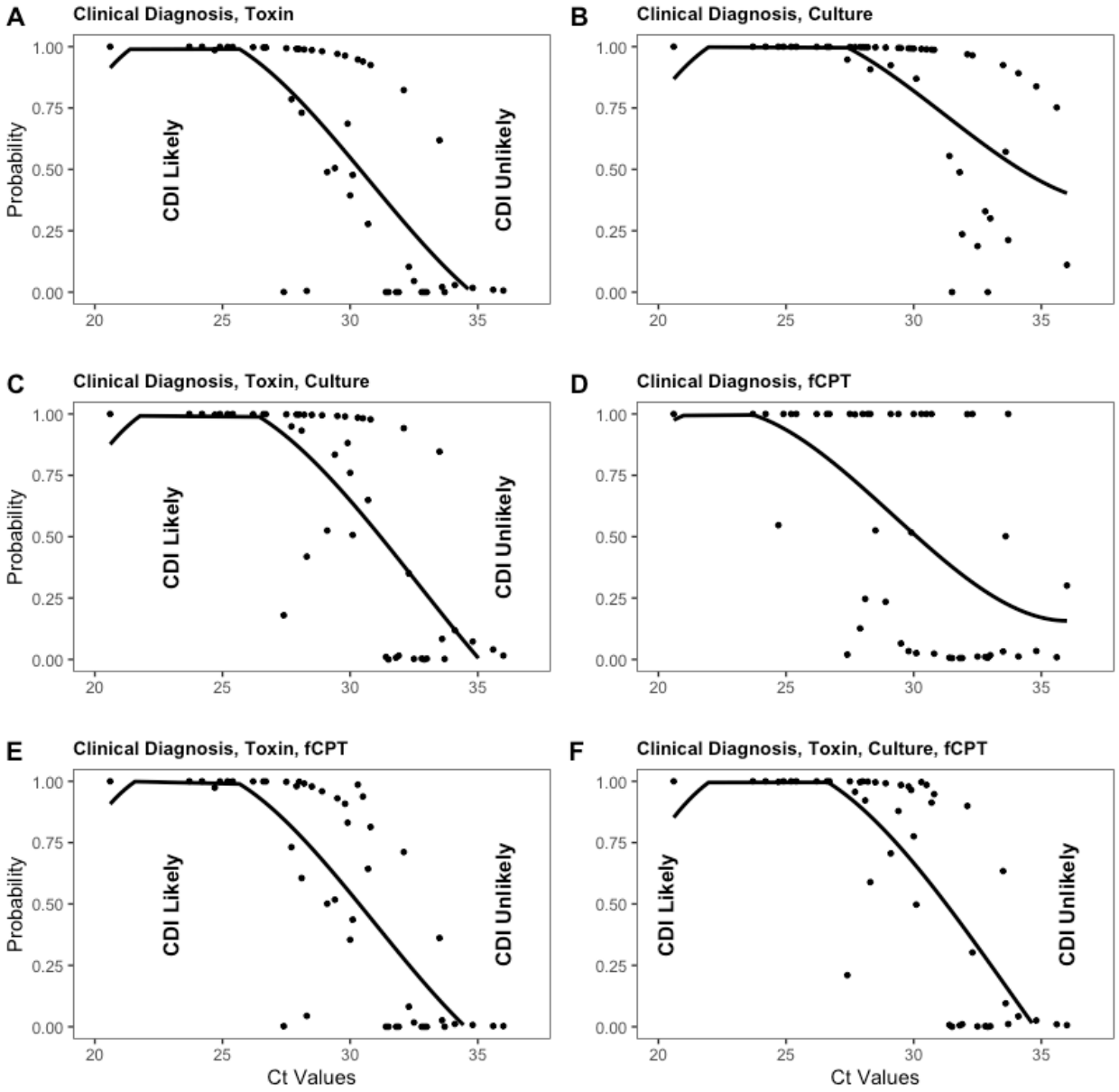
Appendix Figure 2.2: LCA analyses results depicting probability of CDI in relation to the Ct value for Diasorin Simplexa® qPCR. Panels A-F represent each variable: clinical diagnosis, toxin detection by DiaSorin Liaison® XL *C. difficile* Toxin A/B [Toxin], toxigenic culture [Culture], and detection of calprotectin as a maker of inflammation [fCPT].



Appendix Figure 2.3: LCA analyses results depicting probability of CDI in relation to the Ct value for Elite MGB® qPCR. Panels A-F represent each variable: clinical diagnosis, toxin detection by DiaSorin Liaison® XL *C. difficile* Toxin A/B [Toxin], toxigenic culture [Culture], and detection of calprotectin as a maker of inflammation [fCPT].



Appendix Figure 2.4: LCA analyses results depicting probability of CDI in relation to the Ct value for Becton Dickinson BD MAX™ qPCR. Panels A-F represent each variable: clinical diagnosis, toxin detection by DiaSorin Liaison® XL *C. difficile* Toxin A/B [Toxin], toxigenic culture [Culture], and detection of calprotectin as a maker of inflammation [fCPT].



2.8 **References**

1. Centers for Disease Control and Prevention. Antibiotic resistance threats in the United States. Available at: https://www.cdc.gov/drugresistance/biggest-threats.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fdrugresistance%2Fbiggest_threats.html#cdiff. Accessed 27 October 2020.
2. Peng Z, Ling LF, Stratton CW, et al. Advances in the diagnosis and treatment of *Clostridium difficile* infections. Emerg Microbes Infect 2018; 7:15.
3. Daniels LM, Kufel WD. Clinical review of *Clostridium difficile* infection: an update on treatment and prevention. Expert Opin Pharmacother 2018; 19:1759–69.
4. Crobach MJT, Vernon JJ, Loo VG, et al. Understanding *Clostridium difficile* colonization. Clin Microbiol Rev 2018; 31:29.
5. Burnham CA, Carroll KC. Diagnosis of *Clostridium difficile* infection: an ongoing conundrum for clinicians and for clinical laboratories. Clin Microbiol Rev 2013; 26:604–30.
6. Bartsch SM, Umscheid CA, Nachamkin I, Hamilton K, Lee BY. Comparing the economic and health benefits of different approaches to diagnosing *Clostridium difficile* infection. Clin Microbiol Infect 2015; 21:77.e1–9.
7. McDonald LC, Gerding DN, Johnson S, et al. Clinical practice guidelines for *Clostridium difficile* infection in adults and children: 2017 update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). Clin Infect Dis 2018; 66:987–94.
8. Guery B, Galperine T, Barbut F. *Clostridioides difficile*: diagnosis and treatments. BMJ 2019; 366:19.
9. Hanania A, Jiang ZD, Smiley C, Lasco T, Garey KW, DuPont HL. Fecal calprotectin

in the diagnosis of *Clostridium difficile* infection. Infect Dis Clin Pract 2016; 24:31–4.

10. Popiel KY, Gheorghe R, Eastmond J, Miller MA. Usefulness of adjunctive fecal calprotectin and serum procalcitonin in individuals positive for *Clostridium difficile* toxin gene by PCR assay. J Clin Microbiol 2015; 53:3667–9.
11. Kim HN, Kim H, Moon HW, Hur M, Yun YM. Toxin positivity and tcdB gene load in broad-spectrum *Clostridium difficile* infection. Infection 2018; 46:113–7.
12. Polage CR, Gyorke CE, Kennedy MA, et al. Overdiagnosis of *Clostridium difficile* infection in the molecular test era. JAMA Intern Med 2015; 175:1792–801.
13. Hitchcock MM, Holubar M, Hogan CA, Tompkins LS, Banaei N. Dual reporting of *Clostridioides difficile* PCR and predicted toxin result based on PCR cycle threshold reduces treatment of toxin-negative patients without increases in adverse outcomes. J Clin Microbiol 2019; 57:11.
14. Kamboj M, Brite J, McMillen T, et al. Potential of real-time PCR threshold cycle (C-T) to predict presence of free toxin and clinically relevant *C. difficile infection* (CDI) in patients with cancer. J Infect 2018; 76:369–75.
15. Dionne LL, Raymond F, Corbeil J, Longtin J, Gervais P, Longtin Y. Correlation between *Clostridium difficile* bacterial load, commercial real-time PCR cycle thresholds, and results of diagnostic tests based on enzyme immunoassay and cell culture cytotoxicity assay. J Clin Microbiol 2013; 51:3624–30.
16. Crobach MJT, Duszenko N, Terveer EM, Verduin CM, Kuijper EJ. Nucleic acid amplification test quantitation as predictor of toxin presence in *Clostridium difficile* infection. J Clin Microbiol 2018; 56:e01316–17.

17. Davies KA, Planche T, Wilcox MH. The predictive value of quantitative nucleic acid amplification detection of *Clostridium difficile* toxin gene for faecal sample toxin status and patient outcome. PloS One 2018; 13:9.

18. Garvey MI, Bradley CW, Wilkinson MAC, Holden E. Can a toxin gene NAAT be used to predict toxin EIA and the severity of *Clostridium difficile* infection? Antimicrob Resist Infect Control 2017; 6:8.

19. Senchyna F, Gaur RL, Gombar S, Truong CY, Schroeder LF, Banaei N. Clostridium difficile PCR cycle threshold predicts free toxin. J Clin Microbiol 2017; 55:2651–60.

20. Jazmati N, Hellmich M, Ličanin B, Plum G, Kaasch AJ. PCR cycle threshold value predicts the course of *Clostridium difficile* infection. Clin Microbiol Infect 2016; 22:e7–8.

21. Reigadas E, Alcalá L, Valerio M, Marín M, Martin A, Bouza E. Toxin B PCR cycle threshold as a predictor of poor outcome of *Clostridium difficile* infection: a derivation and validation cohort study. J Antimicrob Chemother 2016; 71:1380–5.

22. Halligan S, Altman DG, Mallett S. Disadvantages of using the area under the receiver operating characteristic curve to assess imaging tests: a discussion and proposal for an alternative approach. Eur Radiol 2015; 25:932–9.

23. Origuen J, Orellana MA, Fernandez-Ruiz M, et al. Toxin B PCR amplification cycle threshold adds little to clinical variables for predicting outcomes in *Clostridium difficile* infection: a retrospective cohort study. J Clin Microbiol 2019; 57:e01125–18.

24. Sandlund J, Wilcox MH. Ultrasensitive detection of *Clostridium difficile* toxins

reveals suboptimal accuracy of toxin gene cycle thresholds for toxin predictions. J Clin Microbiol 2019; 57:5.

25. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KG, de Groot JA. Latent class models in diagnostic studies when there is no reference standard—a systematic review. Am J Epidemiol 2014; 179:423–31.

26. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. Am J Epidemiol 1995; 141:263–72.

27. Weakliem DL. A critique of the Bayesian information criterion for model selection. Sociol Methods Res 1999; 27:359–97.

28. Visser I, Speekenbrink M. depmixS4: an R package for Hidden Markov models. J Stat Softw 2010; 36:1–21.

29. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2019.

30. Mweu MM, Wambua J, Njuga F, Bejon P, Mwanga D. Bayesian evaluation of the performance of three diagnostic tests for Plasmodium falciparum infection in a low-transmission setting in Kilifi County, Kenya. Wellcome Open Res 2019; 4:67.

31. Stout JE, Wu Y, Ho CS, et al; Tuberculosis Epidemiologic Studies Consortium. Evaluating latent tuberculosis infection diagnostics using latent class analysis. Thorax 2018; 73:1062–70.

32. Clements MN, Corstjens P, Binder S, et al. Latent class analysis to evaluate performance of point-of-care CCA for low-intensity Schistosoma mansoni infections in

Burundi. Parasit Vectors 2018; 11:13.

33. Jespers V, Menten J, Smet H, et al. Quantification of bacterial species of the vaginal microbiome in different groups of women, using nucleic acid amplification tests. BMC Microbiol 2012; 12:10.

34. Sigera PC, Amarasekara R, Rodrigo C, et al. Risk prediction for severe disease and better diagnostic accuracy in early dengue infection; the Colombo dengue study. BMC Infect Dis 2019; 19:8.

35. Castro-Sesquen YE, Saldana A, Patino Nava D, et al. Use of a latent class analysis in the diagnosis of chronic Chagas disease in the Washington metropolitan area. Clin Infect Dis 2020. Doi:10.1093/cid/ciaa1101

36. Baughman AL, Bisgard KM, Cortese MM, Thompson WW, Sanden GN, Strebel PM. Utility of composite reference standards and latent class analysis in evaluating the clinical accuracy of diagnostic tests for pertussis. Clin Vaccine Immunol 2008; 15:106–14.

3.0 Prospective cohort study evaluating *Clostridioides difficile* qPCR cycle threshold correlation to clinical variables

3.1 Abstract

Background. *Clostridiodes difficile* infections (CDI) is a common burden of disease in acute care facilities and lacks a gold standard clinical diagnostic test to determine CDI. A previous method of latent class analysis established that zones defined by *C. difficile* quantitative polymerase chain reaction (qPCR) cycle threshold (Ct) values can correlate to probability of disease. This paper intends to investigate if these probability zones or Ct values have further correlations beyond diagnostics to clinical variables and indicators of severe CDI.

Methods. A prospective cohort study was conducted with suspected CDI hospital patients. Primary diagnosis, demographic data and indicators of severe disease were captured during CDI including white blood cell count, creatinine, albumin, and C-reactive protein, mortality, and hospital length of stay. Data was evaluated for correlations between captured variables and the patient's qPCR Ct or the previous latent class analysis established Ct defined probability of disease zone: CDI likely (>90% probability), CDI equivocal (<90% and >10%), CDI unlikely (<10%), and negative.

Results. No significant correlations were found between clinical variables and the Ct values or zones.

Conclusions. Ct defined probability zones have been previously shown to assist in diagnostic probability but in this study did not extend useful correlations to other clinical variables.

3.2 Introduction

Clostridioides difficile is one of the most common causes of health care-associated infections and a major clinical burden on acute care facilities [1]. Despite its common prevalence, current technology and practices fall short of providing reliable and rapid laboratory tests that can accurately distinguish true *C. difficile* infections (CDI) from symptomatic colonized patients [2, 3]. Given that CDI is often clinically indistinguishable from other diarrheal illnesses, reliable diagnostic tests are essential for proper clinical management and infection control [3, 4].

Current detection methods include the following diagnostics. Each possesses shortcomings that prevent them from being a true gold standard test for CDI. Cell cytotoxicity neutralization assays (CCNAs), can detect minimal amounts of toxin and organism, but due to dependence on cell cultures, required laboratory expertise, and long turnaround time (>1-2 days), are rarely used as a clinical diagnostic test [2, 5, 6]. The toxin produced by toxigenic *C. difficile* is indicative of CDI [4]. Enzyme immunoassays (EIAs) detecting toxin lack adequate analytical sensitivity to be used as a reliable stand-alone clinical test [4, 7]. Nucleic acid amplification tests (NAAT), such as quantitative polymerase chain reaction (qPCR), targeting conserved regions of toxin-producing genes, are widely used with high analytical sensitivity for detecting the C difficile bacilli, but struggle to distinguish patients colonized by *C. difficile* from clinical disease [5]. The detection of, glutamate dehydrogenase (GDH), produced by both nontoxigenic and toxigenic strains, can provide a rapid and inexpensive screening option with high negative predictive ability. GDH testing must be followed by a confirmatory EIA or NAAT to confirm presence of toxigenic *C. difficile* [2, 7].

Given the trade-offs presented by available diagnostics, clinical guidelines generally recommend multi-step testing algorithms. Still, many laboratories continue to utilize a single-step NAAT such as qPCR as their primary test for CDI [8]. Previous literature has examined in depth the issues of qPCR overcalling disease and investigated how qPCR cycle threshold (Ct) values correlating to disease severity or clinical variables. Ct values have been correlated with the following variables: *C. difficile* bacteria load, fecal toxin, disease severity or variables indicative of disease severity, and CDI clinical outcome variables [9-17]. To address qPCR's prolific use, lack of clear gold standard for clinical disease, and the associated overcalling of disease due to detection of colonized patients, Doolan et al employed latent class analysis (LCA) to correlate qPCR Ct values to probability of CDI [18]. LCA can be utilized to combine multiple variables or results into a single evaluation when a diagnostic lacks a clear gold standard for comparison [19, 20]. Doolan et al.'s LCA incorporated toxin presence, toxigenic culture, and clinical judgement into a latent class model correlating probability of CDI to qPCR Ct values for four commercially available qPCR platforms [18]. For the platforms evaluated, the LCA results identified three zones of Ct values determining the probability of disease in each zone: CDI likely (>90% probability of CDI), CDI equivocal (<90% and >10%), CDI unlikely (<10%). Relevant to this study, the GeneXpert *C. difficile* (Cepheid, Sunnyvale, CA) qPCR Ct zones were separated at Ct 24.00 and Ct 33.61 [18]. The previous LCA evaluation stopped at the attempt to predict true disease status, but can these Ct zones provide additional information on clinical variables that are indicators of CDI severity?

Clinical variables can be used as metrics to determine patient CDI disease severity or indicate likelihood of poor clinical outcomes. These variables include the following: hospital length of

stay, white blood cell (WBC), creatinine, albumin, C-reactive protein (CRP), and mortality. Length of stay has been associated with higher levels of toxin and severe CDI [6]. The Infectious Disease Society of America and Society for Healthcare Epidemiology of America define severe CDI with two metrics, a white blood cell (WBC) count >15000 cells/ml or a serum creatinine >1.5mg/dl [22]. WBC and creatine are well established in the literature for correlating to severity of disease and unfavourable CDI outcomes [22, 23]. Low albumin concentrations are a risk factor of severe CDI and serum albumin is thought to act as a component in the defence mechanism against *C. difficile* [6, 24]. CRP has been identified as a risk factor for severe CDI and an indicator of associated inflammation [23, 25]. These correlations however are not universally accepted as some studies have found no correlation with severity variables to CDI diagnostic results or Ct values [6, 21].

Doolan et al.'s previous retrospective cohort study, Latent Class Analysis for the Diagnosis of *Clostridioides difficile* Infection, demonstrated the utility of Ct threshold to predict a patient's probability of CDI [18]. This publication seeks to determine if the previously established Ct zones provide additional predictive value beyond probability of clinical disease and investigates the correlation of clinical variable and severity metrics with Ct zones and Ct values.

3.3 Methods

Study design, patient population, and ethics

A prospective cohort study was conducted at Foothills Medical Center (FMC), a 1087-bed acute care hospital located in in Alberta, Canada. This study acquired laboratory results from the centralized laboratory services operated by Alberta Precision Laboratories (APL). Data on

suspected CDI patients was collected from September 2018 to May 2020 from inpatient units at FMC. Stool samples were collected using the clinical pipeline at FMC and received by APL. Stool samples were screened using either the C. DIFF QUIK CHEK COMPLETE® (Techlab Inc., Radford, Va) or the Liaison *C. difficile* GDH (DiaSorin, Cypress, CA). Sample negative upon screening were clinically reported immediately, and all positive results were confirmed by GeneXpert *C. difficile* (Cepheid, Sunnyvale, CA) real-time PCR detecting tcdB gene presence. Negative patients included in this analysis were those confirmed negative by PCR confirmation. Additional patient severity and clinical data was extracted from hospital medical records. Ethical approval was obtained through University of Calgary's Institutional Research Information Services Solution (REB-16–0397). Information was protected as per Alberta Health Services protocol and data was de-identified prior to analysis.

Clinical and laboratory data

Patients with suspected CDI were identified (N=9626). CDI suspected patients were defined as patients with a *C. difficile* diagnostic order and stool sample sent for laboratory confirmation. For patients that were positive upon screening and had samples sent for qPCR, Ct values were recorded (N=414). Data for the following indicators of CDI severity was collected for CDI suspected patients during their time in hospital with multiple values per patient (N=149,449 laboratory values): white blood cell (WBC), creatinine, albumin, and CRP. For any one patient if multiple clinical values were present for WBC, creatinine, and CRP, the highest recorded value was used in a 7-day window surrounding time of CDI diagnostic request. For albumin, the lowest clinical value was used in this 7-day window. An additional data extraction was conducted from the hospital's medical record system containing length of stay and primary

diagnosis for suspected CDI patients (N=23,243 records). These data sets were combined to create a final study population for analysis consisting of 339 patients (Figure 3.1). Potential confounders of age and sex were recorded in addition to the presence of a gene target indicative of the *C. difficile* NAP1 strain (confirmed by qPCR). NAP1 is a prevalent *C. difficile* strain that has been found to be correlated with severe disease, poor outcomes, and death [26].

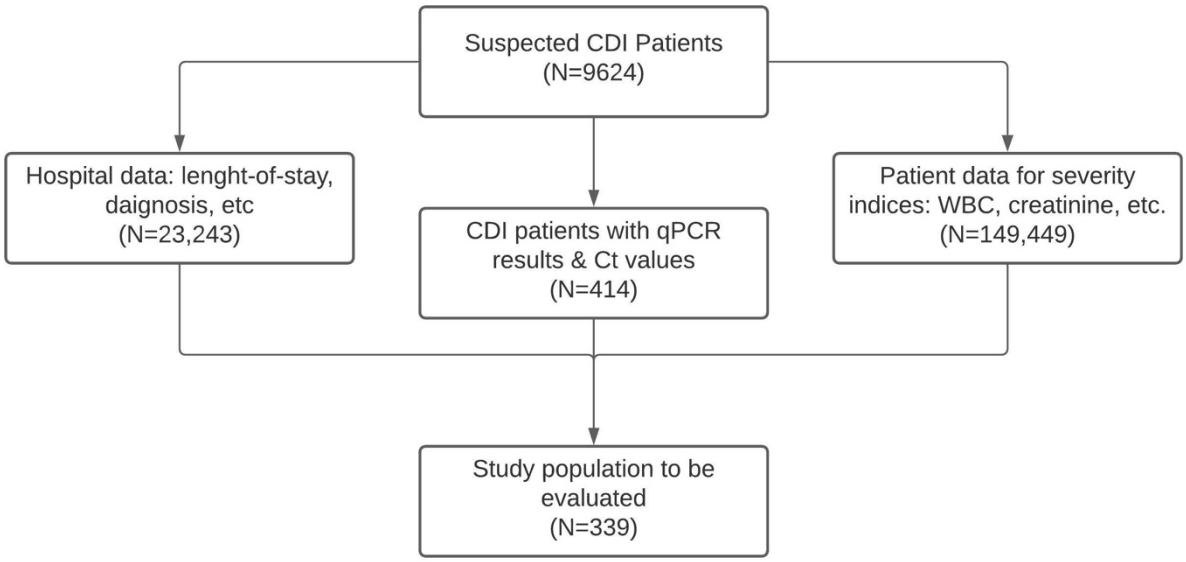


Figure 3.1: Data extraction from hospital (FMC) and laboratory (APL) medical records combined for final study population for analysis.

Statistical analysis

To determine if the Ct value zones as determined by Doolan et al. have correlation beyond clinical diagnosis, patients were first sorted into groups based on CDI qPCR results with associated Ct values determining zone membership, i.e. probability of disease grouping: CDI likely (>90% probability), CDI equivocal (<90% and >10%), CDI unlikely (<10%), and negative [18]. Cut-offs between zones were defined by the GeneXpert qPCR platform's Ct values implemented in the clinical pipeline and values separating the 3 CDI probability zones

correspond to Ct 24.00 and Ct 33.61 [18]. Negative patients were assigned a Ct value of 40 for analysis [18].

Ordinal logistic regression analysis was employed to evaluate the relationship between patients grouped in Ct defined disease probability zones and negative patients to the clinical variable of interest. Zones including negative patients, were set as the dependent variables. Independent variables were WBC, creatinine, CRP, albumin, in hospital mortality, sex, age, presence of NAP1 strain, and hospital length of stay. To broaden the applicability of this analysis beyond Doolan et al's defined zones, a second multiple linear regression analysis was conducted with the same independent variables exchanging zones as the dependent variable for the continuous variable of Ct values.

The study population contained a broad cross section of inpatients at an acute care hospital. These patients suffer from a variety of diagnosis in addition to CDI. Many of these other diagnoses can dramatically affect the clinical variables being examined in this study. We attempted to focus an additional analysis on patients where CDI may play a more major role in influencing the clinical and severity variables of interest. A sub analysis was conducted examining a population of patients identified based on their primary diagnosis. This was a subjective grouping with an intention to limit variations within clinical variables due to other clinical disorders or disease and isolate possible CDI effects. Clinical judgement was used to exclude diagnosis such as infections or renal failure that are known to dramatically affect the chosen clinical variables also influenced by CDI severity. The primary diagnosis indicated as CDI was grouped with other diagnosis as seen in Appendix Table 3.1. The ordinal logistic

regression and the multiple linear regression previously described was repeated using this subset of patients (N=161).

A two-sided alpha of less than 0.05 was considered statistically significant. All analysis was conducted in R-Studio (version 1.4.1103; R Foundation for Statistical Computing, Vienna, Austria) [27].

3.4 Results

Patient Population

After compiling patient data (Figure 3.1), 339 patients were included in this analysis. Patients were 47.2% female with an average age of 61.4 years (18-98 years). The mortality rate was 10.0% during the study, and 4.1% were NAP1 positive. This cohort contained 33.3% qPCR (GeneXpert) negative patients and based on Doolan et al. Ct value zones; 9.7% of patients were identified as unlikely CDI, 47.2% identified as equivocal, and 20.6% identified as likely CDI (Table 3.1). The distribution of values for WBC, creatinine, albumin, and CRP can be graphically visualized based on Ct zone classification of patient in Table 3.2.

Table 3.1: Patient characteristics and Ct value identified zone membership.

Variables	Percentage	Mean	Range
Age (years)	…	61.4	18 to 98
Female	47.2	…	…
WBC (cells/µl)	…	10.2	0.8 to 86.1
Creatinine (mg/dl)	…	2.2	0.27 to 11.24
Length of stay (days)	…	41.0	1 to 1582
CRP (mg/l, N = 187)	…	68.8	1.3 to 437.5
Albumin (g/l, N = 257)	…	23.4	10 to 39
Mortality	10.0	…	…
NAP1 confirmed	4.1	…	…
GeneXpert qPCR (-)	23.0	…	…
Ct Values	…	30.4	19.6 to 40
Zone: Unlikely	9.7	…	…
Zone: Equivocal	47.2	…	…
Zone: Likely	20.6	…	…

N = 339 unless otherwise specified

Abreviations: CDI, Clostridioides difficile infection; -, negative; WBC, White Blood Cell

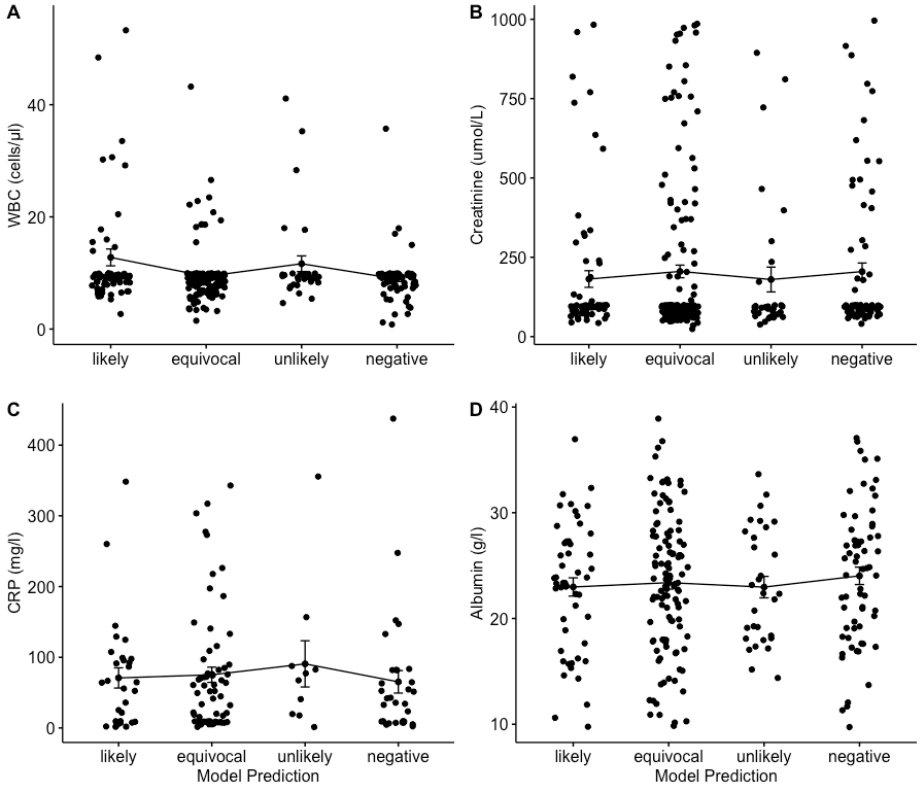


Figure 3.2: Clinical variable WBC, Creatinine, CRP, Albumin distributions grouped by previously established Ct zones (N=339). Abbreviations: WBC, White blood Cells; CRP, C-reactive Protein.

Regression Analyses: All Patients (N=339)

Of the 9 independent variables included in the ordinal logistic regression, 0 were found to be associated with zone membership (Table 3.2). The model was found to have a Akaike information criterion (AIC) of 301.45. For the multiple linear regression analysis with Ct as the dependent variable, 1 variable was found to be associated with Ct values (p-value= 0.04). The multiple linear regression found a significant relationship between age and Ct value showing a 0.077 decrease in age (years) as Ct value increases by one. However, when the p-value of this model based on the global F-statistics was observed, it is non-significant (p-value= 0.14) (Table 3.3).

Table 3.2: Results from ordinal logistic regression conducted on entire study patient population. Dependent variable: Ct Zones

Ordinal Logistic Regression: Patient Population			
Dependent: Ct Zones (Likely, Equivocal, Unlikely, Negative)			
	Coefficients	**Std. Error**	**p-value**
WBC	-0.046	0.03	0.10
Creatinine	0.000	0.00	0.83
CRP	0.000	0.00	0.93
Albumin	-0.023	0.03	0.49
Mortality	-0.579	0.65	0.38
Sex	-0.013	0.37	0.97
Age	-0.020	0.01	0.06
NAP1	-1.247	0.77	0.11
Length of Stay	0.000	0.00	0.77
N= 339, AIC = 301.45			

Table 3.3: Results from multiple linear regression conducted on entire study patient population. Dependent variable: Ct values.

Multiple Linear Regression: Patient Population			
Dependent: Ct Values			
	Coefficients	**Std. Error**	**p-value**
WBC	-0.090	0.07	0.20
Creatinine	0.001	0.00	0.76
CRP	0.000	0.01	0.99
Albumin	-0.091	0.12	0.44
Mortality	-2.467	2.28	0.28
Sex	-0.848	1.30	0.52
Age	-0.077	0.04	**0.04**
NAP1	-4.345	2.79	0.12
Length of Stay	0.001	0.00	0.85
N= 339, Adjusted R-squared= 0.04, **p-value= 0.14**			

Sub-analysis based on primary diagnosis (N=161)

With a subset of the population is being evaluated, the sample size decreased to 161 patients and the same two regression as previously described were repeated. Of the 9 independent variables included in the ordinal logistic regression, 1 variable, age, was found to be associated with zone membership (p-value= <0.01, Table 3.4). The AIC improved in this analysis with a decrease to 117.9. Using the same patient subset in the multiple linear regression analysis with Ct as the dependent variable, 1 variable, age, again was found to be associated with Ct values (p-value= 0.01, Table 3.5). The global p-value of the multiple linear regression model is non-significant (p-value= 0.29).

Table 3.4: Results from ordinal logistic regression conducted on subset of study patient population. Dependent variable: Ct Zones

Ordinal Logistic Regression: Patient Population Subset			
Dependent: Ct Zones (Likely, Equivocal, Unlikely, Negative)			
	Coefficients	**Std. Error**	**p-value**
WBC	-0.051	0.08	0.52
Creatinine	0.000	0.00	0.97
CRP	-0.003	0.01	0.60
Albumin	0.011	0.06	0.85
Mortality	0.135	1.49	0.93
Sex	0.891	0.68	0.19
Age	-0.063	0.02	**<0.01**
NAP1	-2.276	1.22	0.06
Length of Stay	0.000	0.00	0.87
N= 161, AIC = 117.9			

Table 3.5: Results from multiple linear regression conducted on entire study patient population. Dependent variable: Ct values.

Multiple Linear Regression: Patient Population Subset			
Dependent: Ct Values			
	Coefficients	**Std. Error**	**p-value**
WBC	-0.080	0.25	0.75
Creatinine	0.002	0.00	0.63
CRP	-0.003	0.02	0.84
Albumin	0.045	0.20	0.82
Mortality	-0.391	4.77	0.94
Sex	1.638	2.16	0.45
Age	-0.177	0.07	**0.01**
NAP1	-7.355	3.76	0.06
Length of Stay	0.000	0.01	0.94
N= 161, Adjusted R-squared= 0.06, **p-value= 0.29**			

3.5 Discussion

The results of the analysis found minimal correlations between Ct values or Ct zones and the nine independent clinical variables. The significant correlation found between Ct and age is contradicted by non-significant p-values of the models based on global F-statistics. Given the F-statistic findings and considering the quantity of variables being evaluated over multiple regressions, it is suggestive that these significant finding regarding Ct and age are likely a function of family-wise type I error. Regardless, when examining the models' coefficient values for age, they indicate a rate of change per unit Ct value that is minimal and could be interpreted as clinically non-relevant. Given these interpretations, we did not consider age to be a meaningful significant finding of this analysis.

In a prior publication, Doolan et al. demonstrated the potential utility of Ct values to render the probability of clinical CDI and its potential usefulness for implementation in diagnostic strategy [18]. The analyses in this paper examined if the Ct values and zones further correlated with variables and known indicators of severe disease or poor outcomes. Interestingly, the analysis did not support that severity variables or length of stay correlated to Ct. These findings add to a discussion that correlations of clinical variables to CDI qPCR Ct values is not straightforward. Publications such as Reigadas et at do find correlations with clinical variables [17]. While other works such as Origuen et al or Bosch et al, do not support that Ct values add value to predictive models for severity variables [28, 29]. The findings of this paper agree with the later in that these correlations are not observed. Additional investigation is still needed to define when meaningful correlations do and do not exist and when these correlations can provide information to inform

clinical practices. Differences in finding could likely be a function of the patient populations being investigated.

This study was conducted using a broad cohort of patients with widely varying primary diagnoses across hospital inpatient wards. Patients in hospital for diagnoses other than CDI, that strongly affect clinical variables, introduce variations that are unrelated to CDI and thus would not be correlated to *C. difficile* qPCR Ct. The primary investigation of this paper analyzed the entire study population with a sub analysis isolated a population where the observed differences in variables might be more attributed to CDI. Given that neither of these approaches provided a correlation of interest, the possibility of type II error and sufficient power to detect differences between Ct values should be evaluated. If we examine the power of the primary analysis with 339 patients, 9 predictor variables, and a relatively small effect size of 0.05, we observe a sufficient power of 83.9%. When the sub analysis was conducted, the ordinal logistic regression did see a decrease in AIC indicative of a decrease in prediction error but this was accompanied by a dramatic decrease in sample size to 161 patients. With this reduction of the sample size, a similar moderately small effect size, and the same number of predictor variables, the power of the analysis was reduced to an insufficient 46.1%. With this reduction of power, it is possible that the null hypothesis was incorrectly accepted in the sub analysis and that a correlation might yet be undetected. Future studies should be specifically designed and powered to investigate subpopulations of CDI patients where CDI is the primary contributor to clinical symptoms while controlling for clinical comorbidities and interfering diagnosis.

An additional potential limitation of this study was the filtering of patient data for highest value of WBC, lowest albumin, etc. in a 7-day window. The intention was to identify moments of potential severe disease in a window surrounding the CDI diagnostics request. It is possible that this method could have introduced bias and collected data that represented the population as more severe than was factual. Future studies should closely consider data collection choices and incorporate in-depth chart review in order to identify a more exacting method of isolating patient data prior to analysis.

Future analysis could also expand to include examining the change of clinical variables and Ct values over time and the potential relationships. These observations could provide further insight into positive or negative disease progression and correlation to Ct values during different stages of disease. Additional patient data including but not limited to presence of colectomy, movement to and from ICU, and how metrics such as creatine change as patient severity changes might illuminate further possible correlations or predictions.

It is possible that the finding of this work and others are indicating that Ct values are just not the best way to correlate a laboratory available value with metrics of CDI disease severity. Ct value correlations are based on several major assumptions. Primarily that Ct value are correlated with bacterial load/burden and toxin presence. Toxin is a critical component for the progression of CDI pathophysiology leading to clinical symptoms, and thus these correlations have potential to further correlate to disease severity [9, 13, 14]. These assumptions however are not universally observed. For instance, Pollock et al using ultrasensitive quantitative immunoassays to quantify toxin in stool did not find a significant difference in toxin load between asymptomatic carriers of

C. difficile (exhibiting no clinical symptoms) and symptomatic patients confirmed with CDI [30]. Indicators of how a patient's immune system is reacting to the toxin presence could be a better predictor of disease severity and associated variables. This is supported by both Feghaly et al demonstrating CDI outcomes were correlated with levels of inflammatory cytokines but not bacterial burden, and further work by Kelly et al demonstrating how immune markers can distinguish between CDI and colonization [31, 32]. The positive and at times contradicting findings in literature spanning both Ct and inflammatory marker-based studies support the idea that multiple factors must be aligned to have clinical disease caused by *C. difficile*. The presence of *C. difficile*, toxin load, and a host immune response are all necessary for CDI but appear insufficient to cause disease alone. The next steps in CDI disease investigations will be to continue to understand the interactions and roles that each factor plays in the disease process and how this knowledge can be used to improve diagnostics and patient outcomes.

Conclusion

The intention of this publication's was to examine a broad base of patients at a single time point investigating Ct correlations across the breadth of the inpatient population and diagnosis. This study specifically focused on results from samples collected from suspected CDI patients. The overarching goal was to determine if with a diagnostic test requested by a clinician and the associated Ct values generated, can we provide additional valuable information for CDI clinical management. Unfortunately, future studies are needed to fully understand this question. This paper does not detract from the previous publication by Doolan et al establishing Ct zone's ability to predict probability of CDI, it instead helps to further defines the Ct model and the previous LCA model's current limitation to disease prediction. This work provides evidence

supporting the need for further investigation on the correlations of clinical variables to laboratory tests such as qPCR Ct values in *C. difficile* patients.

3.6 Chapter 3 Appendix Material

Appendix Table 3.1: Grouping of primary diagnostics identifying patients where CDI would more likely affect clinical variables.

Category	Primary Diagnosis
Evaluated in secondary analysis (N=161)	Ac subendocardial myocardial infarction
	Ac transv myelitis in demyelin dis CNS
	Acute transmural MI of oth site
	Alzheimer's disease unspecified
	Anoxic brain damage NEC
	Aortic (valve) stenosis
	Ath hrt dis native coron art
	Atrioventricular block complete
	Benign paroxysmal vertigo
	Beriberi
	Bronchopneumonia unspecified
	Calc gallblad w ac cholecyst w/o obstrct
	Cardiogenic shock
	Care inv use of rehab procedure NOS
	Central cord lesion cervical spinal cord
	Cereb infarct dt embolism cerebral art
	Cereb infarct dt embolism precereb art
	Cereb infarct dt occlusion cereb art NOS
	Chronic or unspec duodenal ulcer w haem
	Chronic or unspec gastric ulcer w haem
	Congestive heart failure
	COPD with acute exacerbation unspecified
	COPD with acute lower resp infection
	Coxarthrosis unspecified
	Decub ulcer & press area unspec
	Decub ulcer & press area unspec
	Delirium superimposed on dementia
	Diaph hernia without obs or gangrene
	Diffuse brain inj w/o open intracran wd
	Dissection of aorta
	Enterocolitis dt clostridium difficile
	Fistula of vagina to large intestine
	Fluid overload

Focal brain injury w/o open intracran wd
Fx bone follw insert orth implnt unspec
Fx of second cervical vertebra, closed
Fx of thoracic vertebra T1 - T6, closed
GB lacr w/o BD inj w/o opn wnd cav
Hypertensive heart disease
Hypotension unspecified
Inflammatory disorders seminal vesicle
Intertrochanteric fracture, closed
Intestinal adhesions with obstruction
Intracereb haem in hemisphere cortical
Intracerebral haem intraventricular
Intracranial injury unspec
Iron deficiency anaemia dt blood loss
Lumbar and oth I/V disc disrd w radiclpth
Malaise and fatigue
Mech comp GI prosth dev impl gft
Mech comp oth spec int prosth dev impl
Ment/beh disrd dt alco use withdrawal st
Mild cognitive disorder
Multiple fractures of 2 - 4 ribs, closed
NA
Nausea with vomiting
Noninfect gastroenteritis and colitis NOS
Obstruction of bile duct
Orthostatic hypotension
Oth spec crystal arthropathies lower leg
Other & unspec convulsions
Other and unspecified dysphagia
Other delirium
Other forms of acute pericarditis
Other fracture of femoral neck, closed
Other giant cell arteritis
Other hydrocephalus
Other impaction of intestine
Other spec diseases biliary tract
Pain management planning
Palliative care
Pleural effusion NEC
Pneumonia unspecified
Pneumonitis due to food and vomit
Poisoning by cocaine
Poisoning by fentanyl and derivatives
Poisoning by hydromorphone
Postproc pelvic peritoneal adhesions
Preparatory care for subsequent Rx NEC
Psn oth parasympatholytic spasmlytic NEC
Resp failure unspec type 1 [hypoxic]

	Sick sinus syndrome
	Status epilepticus unspecified
	Subdural haem (acute)(nontraumatic)
	Traumatic ischaemia of muscle
	Traumatic subdural haemorrhage
	Type 2 DM with foot ulcer
	Ulcer low limb not elsewhere classified
	Unspec fx low (distal) end of femur open
	Unspecified dementia
	Varicose veins legs w ulcer and inflam
	Vascular comp following a procedure NEC
	Vascular myelopathies
	Ventricular fibrillation
	Ventricular tachycardia
	Volvulus
Excluded in secondary analysis (N=178)	Abscess of vulva
	Acute appendcts w local peritonitis
	Acute cholecystitis
	Acute leukaemia of unspecified cell type
	Acute lymphoblastic leukaemia [ALL]
	Acute myeloblastic leukaemia
	Acute pancreatitis, unspecified
	Acute prostatitis
	Acute renal failure unspecified
	Acute resp failure type 2 [hypercapnic]
	Acute vascular disorders of intestine
	Adult respiratory distress syndrome
	Alcoholic cirrhosis of liver
	Alcoholic hepatic failure
	Alcoholic hepatitis
	Bacterial infection unspecified
	Benign neoplasm of lower jawbone
	Cachexia
	Calculus of kidney
	Chemotherapy session for neoplasm
	Chronic cholecystitis
	Chronic kidney disease stage 5
	Chronic myelomonocytic leukaemia
	Crohn's disease of large intestine
	Crohn's disease of small intestine
	Crohn's disease unspecified
	Cutan abs furuncle carbuncle other sites
	Cutan abscess furuncle and carbuncle trunk
	Diffuse large B-cell lymphoma
	Diverticlr dis large intest w perf and abs
	Diverticlr dis lrg intest w/o perf and abs
	Enterostomy malfunction, NEC
	Fever unspecified
	Follicular lymphoma grade II
	Gastric ulcer acute with haemorrhage
	Gastro-oesophageal laceration-haem syndr
	Gastrointestinal haemorrhage NOS

Graft-versus-host reaction or disease
Haemoptysis
Heart transplant rejection
Hyperosmolality and hypernatraemia
Infect and infl reaction dt knee prosth
Infection foll infus transfn thrpc injct
Infection foll infus transfn thrpc injct
Infection following a procedure NEC
Infection of below knee amputation stump
Infectn inf dt oth int prosth dev impl
Infectn intervertebral disc lumbosacral
Influenza w pneum ident seasonal virus
Influenza w pneum ident seasonal virus
Kidney transplant failure
Kidney transplant rejection
Left sided colitis
Local infectn inf dt central ven cath
Localized enlarged lymph nodes
Malgt neoplm ovary u/bilateral not spec
Malignant lesion oesophagus unspecified
Malignant neoplasm of anal canal
Malignant neoplasm of appendix
Malignant neoplasm of bladder unspec
Malignant neoplasm of cardia
Malignant neoplasm of endometrium
Malignant neoplasm of ileum
Malignant neoplasm of lower limb
Malignant neoplasm of prostate
Malignant neoplasm rectosigmoid junction
Mantle cell lymphoma
Meningitis unspecified
Multiple myeloma
Myelodysplastic syndrome unspecified
Neutropenia
Oth and unsp gastroe and colitis inf origin
Other acute renal failure
Other agranulocytosis
Other bacterial infections of site NOS
Other Crohn's disease
Other spec sepsis
Pneumocystosis
Pyonephrosis
Rotaviral enteritis
Sec malgt neoplasm bone and bone marrow
Sec malgt neoplm brain cerebral meninges
Sec malgt neoplm lrg intestine and rectum
Secondary malignant neoplasm right lung
Sepsis due to anaerobes
Sepsis due to E. coli
Sepsis due to enterococcus
Sepsis unspec
Staphylococcal infection unspec site
Streptococcal sepsis unspecified
Toxic gastroenteritis and colitis
Tubulo-interstitial nephritis NOS

	Type 1 DM w establish adv renal disease
	Type 1 DM with keto & lactic acidosis
	Type 1 DM with ketoacidosis
	Type 2 DM w establish adv renal disease
	Type 2 DM w foot ulcer w gangrene
	Type 2 DM w periph angiopathy gangr
	Type 2 DM with ketoacidosis
	Umbilical hernia with obs w/o gangrene
	Unspec infectn inf dt central ven cath
	Urinary tract infection site not spec
	Viral carditis
	Viral intestinal infection unspecified
	Zoster encephalitis

3.7 References

1. Magill SS, Edwards JR, Bamberg W, et al. Multistate Point- Prevalence Survey of Health Care- Associated Infections. New England Journal of Medicine **2014**; 370(13): 1198-208.
2. Crobach MJT, Planche T, Eckert C, et al. European Society of Clinical Microbiology and Infectious Diseases: update of the diagnostic guidance document for Clostridium difficile infection. Clinical Microbiology and Infection **2016**; 22: S63-S81.
3. Daniels LM, Kufel WD. Clinical review of *Clostridium difficile* infection: an update on treatment and prevention. Expert Opinion on Pharmacotherapy **2018**; 19(16): 1759-69.
4. Bartsch SM, Umscheid CA, Nachamkin I, Hamilton K, Lee BY. Comparing the economic and health benefits of different approaches to diagnosing *Clostridium difficile* infection. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases **2015**; 21(1): 77.e1-9.
5. Crobach MJT, Baktash A, Duszenko N, Kuijper EJ. Diagnostic Guidance for *C. difficile* Infections. In: Mastrantonio P, Rupnik M. Updates on Clostridium Difficile in Europe: Advances in Microbiology, Infectious Diseases and Public Health, Vol 8 Vol. 1050, **2018**:27-44.

6. Planche TD, Davies KA, Coen PG, et al. Differences in outcome according to Clostridium difficile testing method: a prospective multicentre diagnostic validation study of *C difficile* infection. Lancet Infectious Diseases **2013**; 13(11): 936-45.

7. Guery B, Galperine T, Barbut F. *Clostridioides difficile*: diagnosis and treatments. Bmj-British Medical Journal **2019**; 366: 19.

8. Rao KN, Malani PN. Diagnosis and Treatment of *Clostridioides (Clostridium) difficile* Infection in Adults in 2020. Jama-Journal of the American Medical Association **2020**; 323(14): 1403-4.

9. Kamboj M, Brite J, McMillen T, et al. Potential of real-time PCR threshold cycle (C-T) to predict presence of free toxin and clinically relevant *C. difficile* infection (CDI) in patients with cancer. Journal of Infection **2018**; 76(4): 369-75.

10. Dionne LL, Raymond F, Corbeil J, Longtin J, Gervais P, Longtin Y. Correlation between *Clostridium difficile* Bacterial Load, Commercial Real-Time PCR Cycle Thresholds, and Results of Diagnostic Tests Based on Enzyme Immunoassay and Cell Culture Cytotoxicity Assay. Journal of Clinical Microbiology **2013**; 51(11): 3624-30.

11. Kim HN, Kim H, Moon HW, Hur M, Yun YM. Toxin positivity and tcdB gene load in broad-spectrum *Clostridium difficile* infection. Infection **2018**; 46(1): 113-7.

12. Crobach MJT, Duszenko N, Terveer EM, Verduin CM, Kuijper EJ. Nucleic Acid Amplification Test Quantitation as Predictor of Toxin Presence in *Clostridium difficile* Infection. Journal of Clinical Microbiology **2018**; 56(3).

13. Senchyna F, Gaur RL, Gombar S, Truong CY, Schroeder LF, Banaei N. *Clostridium difficile* PCR Cycle Threshold Predicts Free Toxin. Journal of Clinical Microbiology **2017**; 55(9): 2651-60.

14. Davies KA, Planche T, Wilcox MH. The predictive value of quantitative nucleic acid amplification detection of *Clostridium difficile* toxin gene for faecal sample toxin status and patient outcome. Plos One **2018**; 13(12): 9.

15. Garvey MI, Bradley CW, Wilkinson MAC, Holden E. Can a toxin gene NAAT be used to predict toxin EIA and the severity of *Clostridium difficile* infection? Antimicrobial Resistance and Infection Control **2017**; 6: 8.

16. Jazmati N, Hellmich M, Licanin B, Plum G, Kaasch AJ. PCR cycle threshold value predicts the course of *Clostridium difficile* infection. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases **2016**; 22(2): e7-e8.

17. Reigadas E, Alcala L, Valerio M, Marin M, Martin A, Bouza E. Toxin B PCR cycle threshold as a predictor of poor outcome of *Clostridium difficile* infection: a derivation and validation cohort study. Journal of Antimicrobial Chemotherapy **2016**; 71(5): 1380-5.

18. Doolan CP, Louie T, Lata C, et al. Latent class analysis for the diagnosis of *Clostridioides difficile* infection. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **2020**.

19. Joseph L, Gyorkos TW, Coupal L. BAYESIAN-ESTIMATION OF DISEASE PREVALENCE AND THE PARAMETERS OF DIAGNOSTIC-TESTS IN THE ABSENCE OF A GOLD STANDARD. American Journal of Epidemiology **1995**; 141(3): 263-72.

20. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KGM, de Groot JAH. Latent Class Models in Diagnostic Studies When There is No Reference Standard-A Systematic Review. American Journal of Epidemiology **2014**; 179(4): 423-31.

21. Humphries RM, Uslan DZ, Rubin Z. Performance of *Clostridium difficile* Toxin Enzyme Immunoassay and Nucleic Acid Amplification Tests Stratified by Patient Disease Severity. Journal of Clinical Microbiology **2013**; 51(3): 869-73.

22. McDonald LC, Gerding DN, Johnson S, et al. Clinical Practice Guidelines for *Clostridium difficile* Infection in Adults and Children: 2017 Update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). Clinical Infectious Diseases **2018**; 66(7): E1-E48.

23. Abou Chakra CN, Pepin J, Sirard S, Valiquette L. Risk Factors for Recurrence, Complications and Mortality in *Clostridium difficile* Infection: A Systematic Review. Plos One **2014**; 9(6): 14.

24. di Masi A, Leboffe L, Polticelli F, et al. Human Serum Albumin Is an Essential Component of the Host Defense Mechanism Against *Clostridium difficile* Intoxication. Journal of Infectious Diseases **2018**; 218(9): 1424-35.

25. Hardt C, Berns T, Treder W, Dumoulin FL. Univariate and multivariate analysis of risk factors for severe *clostridium difficile*-associated diarrhoea: Importance of co-morbidity and serum C-reactive protein. World Journal of Gastroenterology **2008**; 14(27): 4338-41.

26. See I, Mu Y, Cohen J, et al. NAP1 Strain Type Predicts Outcomes From *Clostridium difficile* Infection. Clinical Infectious Diseases **2014**; 58(10): 1394-400.

27. Team RC. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, **2019**.

28. Origuen J, Orellana MA, Fernandez-Ruiz M, et al. Toxin B PCR Amplification Cycle Threshold Adds Little to Clinical Variables for Predicting Outcomes in *Clostridium*

difficile Infection: a Retrospective Cohort Study. Journal of Clinical Microbiology **2019**; 57(2).

29. Bosch DE, Mathias PC, Krumm N, Bryan A, Fang FC, Greninger AL. Elevated White Blood Cell Count Does Not Predict *Clostridium difficile* Nucleic Acid Testing Results. Clinical Infectious Diseases **2021**.

30. Pollock NR, Banz A, Chen XH, et al. Comparison of *Clostridioides difficile* Stool Toxin Concentrations in Adults With Symptomatic Infection and Asymptomatic Carriage Using an Ultrasensitive Quantitative Immunoassay. Clinical Infectious Diseases **2019**; 68(1): 78-86.

31. El Feghaly RE, Stauber JL, Deych E, Gonzalez C, Tarr PI, Haslam DB. Markers of Intestinal Inflammation, Not Bacterial Burden, Correlate With Clinical Outcomes in *Clostridium difficile* Infection. Clinical Infectious Diseases **2013**; 56(12): 1713-21.

32. Kelly CP, Chen XH, Williams D, et al. Host Immune Markers Distinguish *Clostridioides difficile* Infection From Asymptomatic Carriage and Non-C. difficile Diarrhea. Clinical Infectious Diseases **2020**; 70(6): 1083-93.

4.0 Common Discussion

4.1 Context and limitations

Despite the high prevalence of *C. difficile* and progress in diagnostics, the ability to accurately diagnose clinical CDI remains allusive. Given the prevalence of qPCR diagnostics for CDI diagnostics, this is not the first work to evaluate Ct value correlation to disease or clinical variables. Previous works, relying heavily on ROC analysis, sought for single Ct value cut-point to distinguish clinical disease from colonization and correlate with clinical variables and toxin load. Cycle threshold was critically established to be correlated with free toxin levels as toxin is a component necessary for CDI pathophysiology and indicative of true disease [1-4]. Studies also examined Ct direct correlation to diagnosis of CDI [2, 3]. This work adds significant value to the collective knowledge of CDI diagnostics in several ways. This is the first application of LCA to CDI diagnostics and helps to elucidates why previous attempts to make a predictive model using Ct values have failed. It is a move away from an overly simplistic model of a single Ct cut point and takes into the consideration the more complex biology and clinical presentation involved with CDI. It identified probability zones (CDI likely, CDI equivocal, and CDI unlikely) using latent class analyses and attempts to better define the limits of Ct based models' ability to predict true disease.

This work continues to probe these questions deeper with an investigation into other clinical variables as indicators of severe disease and their relationship with Ct. This works adds to the knowledge that correlations of clinical variables to Ct values is not straightforward. Some papers such as Reigadas et at did find correlations with variables indicative or poor outcome [5]. However, other published works such as Origuen et al or Bosch et al, support contradicting

evidence that Ct values do not add value to predictive models for severity variables [6, 17]. This works adds evidence that these correlations are not clearly defined. More investigations should be made to determine when these correlations with Ct values do and do not exist and when this can provide useful information to inform clinical practices.

The conclusion that the Ct model were not a useful predictor of severity metrics was not entirely surprising due to the contradicting literature and this leads to further questions. The analysis was conducted on a broad cross section of inpatients with CDI and suspected CDI. The models were unable to predict indicators of CDI severity such as WBC, albumin, CRP possibly due to these variables being affected by a plethora of other clinical conditions. This caused a huge variance within the patient population for these variables. Future studies should further investigate not only different subsets of patients, and different permutations of predictor variables, but also investigate other potential predictors beyond Ct that are commonly employed in hospital. Although this work found a negative result, it adds to the collective knowledge and foundation surrounding Ct value-based predictions for CDI and by no means discredits the use of Ct value based predictions to determine the probability of disease. It helps to define the limitation of the model and where Ct values provide value and where more work is needed.

There is an additional avenue to build further upon this work. The models and statistical investigations of CDI in this thesis are limited by the selected input variables. As the collective knowledge around CDI diagnostics continues to grow, these models can be updated to potentially decrease the breadth of area that is defined as an equivocal Ct zone with a hope to provide more detailed information to clinicians. In the latent class analyses, the presence of an equivocal zone does not discredit the potential for clinically relevant implementation. It

alternatively adds value by more clearly defining the "known unknown" of the diagnostic test and Ct associated correlations. Not clearly defining this zone of uncertainty is one of the shortcomings of previous attempts at Ct cut points for diagnostics that this work addresses.

LCA has seen increasing application in medical research in recent years and was chosen as the method for the first study for many of the same reasons driving its growing popularity, the ability to analyze complex data with a lack of clear gold standard and the more readily available computations power of home computers seen over the past several decades [7, 8]. CDI lacks a clear GS test or clinical diagnostic strategy to accurately predict true disease [8-10]. LCA's ability to integrate multiple gold standards into a single probability model is advantageous in predicting a "true" diagnosis [11]. LCA methods are useful in predicting constructs that we cannot directly measure, or latent classes [10]. For instance, in the case of CDI, we can measure a patient's symptoms, if they are colonized by the *C. difficile* bacilli, or if there is toxin present, but we cannot directly measure "CDI disease". LCA additionally and correctly assumes that these observed variables (the model inputs) are imperfect predictors of CDI status (the latent class). This is not the case in more conventional analyses where a gold standard is assumed to be a perfect predictor of disease and only one comparison can occur in each model. Much of previous work described on the correlations of Ct values to CDI employ some form of ROC or gold standard analysis. These overly simplified method relying on a gold standard, do not incorporate the complexity of diagnostics, and provide results that are not reflective of the complexity of the disease process. As a result, disease prediction based on ROC curves or other methods have yet to yield results that can be used in a clinical setting.

One potential weakness of the LCA model is the necessity to assume a state of conditional independence among the observed variables. The assumption is that the observed variables are conditionally independent of the presence or absence of CDI, or to rephrase that the results of one variable or test that contributes to the LCA is independent and gives no information regarding the other tests or variables [8, 9]. This is necessity for LCA but often fails in practice as correlations between tests do exist as they are testing for metrics of the same disease on the same patient. Additionally, the amount of a substance or organism present will give non consistence false negatives when quantities near the LOD of a particular test, i.e., more false-negative values with lower amounts of target [8]. For example, low levels of *C. difficile* toxin in a patient's stool will give more false negatives than a patient with very high levels of toxin. This does not invalidate the models but contributes to their imperfect predictive ability and acknowledges a necessary assumption. LCA is a proven statistical tool that when used properly can provide new information and insight into the value and nuance of a diagnostic tests [10].

4.2 Clinical application

With the current state of this work there is an ability to inform clinical application of Ct based diagnostics for CDI. The clinical diagnosis of CDI using the described LCA model and associated Ct value cut-offs are clearly not a perfect representation of disease status. The model can however allow clinicians to reduce overdiagnosis of CDI by titrating qPCR to better predict "true" CDI. This not only can positively affect the health system, but accurate diagnosis is the best thing for patients.

Clinical implementation would focus on the language of the test's diagnostic result. Instead of a dichotomous qPCR result of positive or negative, a four-tiered result might dramatically affect clinical practice. The language around the results would need to be tested and refined further but an example of result options are as follows: 1) CDI likely recommend CDI treatment, 2) CDI possible rule out other causes, 3) CDI unlikely, 4) Negative. CDI likely patients would have a greater than 90% possibility of disease and laboratory results should encourage treatment. For CDI possible patients, language should encourage the clinician to use clinical judgement to assess treatment options. For CDI unlikely patients with a probability of disease of less than 10%, language should discourage treatment but leave open the possibility to be overruled by clinical judgement. Patients negative by qPCR, NAATs, or other tests with a high negative predictive ability should generally not be considered CDI positive, treated or isolated. The intention of this implementation would be to titrate down over treatment of CDI. Future investigation should examine how strong of language, or where to place probabilities of disease cut-offs to titrate down treatment of *C. difficile* colonization while introducing minimal instances of withholding treatment when it is clinically required. This further investigation would need to balance both the economic and clinical costs of false negative and false positive results and their effects on patients. If an imperfect system with language that introduced some doubt into the clinician mind to encourage deeper clinical diagnosis of patients were introduced, there is potential for an immediate reduction in overtreatment of colonized patients with limited risk to patient wellbeing. It has been demonstrated that up to half of patients confirmed by molecular testing as CDI positive do not experience adverse effects without antibiotic treatment [12].

4.3 Impact

A decrease in overdiagnosis, if implemented broadly, could have far reaching impact, increase targeted treatments, and decrease the overuse of antibiotics used in mistreatment of CDI. The health system here in Calgary uses a two-step algorithm to screen suspected CDI patients but confirms all diagnosis with qPCR. The Ct model corresponding to the Calgary implemented GeneXpert C. difficile (Cepheid, Sunnyvale, CA) qPCR, has Ct zones delineated at Ct 24.00 and 33.61. These Ct values separate the three zones: CDI likely (>90% probability of CDI), CDI equivocal (<90% and >10%), CDI unlikely (<10%). The distribution of patients during the second study (Chapter 2) consisted of 12.64% of Ct values in the CDI likely zone, 66.5% of Ct values in the equivocal zone, and 20.9% of Ct values in the CDI unlikely zone. These results suggest that qPCR is potentially over diagnosing patients in excess of 20.9%. This estimate would grow if a reduction of diagnosis occurred for patients with Ct values falling in the equivocal zone; many patients in this zone will not have CDI. Evaluations of total costs attributed to CDI directly related expenses range from $2,992 to $29,000 [13]. In a retrospective cohort study by Pereira et al in Ontario they found a median cost attributed to CDI was approximately $11,917 [14]. At Foothills Medical Center in Calgary, Alberta there is currently an average of 373 qPCR confirmed positive CDI inpatients annually. If the diagnostic model of this thesis were to be implemented at FMC alone, and we consider that >20.9% of CDI cases could be false positives, using the Ontario CDI cost estimate, this would result in a cost saving of over $929,000 annually. These savings would occur by simply changing how CDI diagnostics are reported with no additional laboratory infrastructure changes and using information already produced at the clinical laboratory (APL). The dramatic savings observed at a single hospital would rapidly increase if expanded to a larger health system. In the United States more than

300,000 annual hospitalizations involve a CDI diagnosis, costing the health system up to $4.9 billion [12]. Implementation of the diagnostics model introduced in this work would result in an estimated savings of over $747 million (assuming continued widespread reliance on qPCR diagnostics, etc.). Overdiagnosis of CDI is not a new concept and this thesis both adds to the argument against CDI overdiagnosis and provides a solution for improving diagnostic accuracy with the current infrastructure [12]. In addition to cost savings, more targeted diagnostics will lead to more targeted care for patients and improved antibiotic stewardship.

4.4 Conclusion

The aim of this thesis was to develop methods to improve CDI diagnostics in health systems with easily implementable techniques. This thesis describes a method to better distinguish clinical CDI disease using existing and implemented diagnostic technology. With the lack of a CDI gold standard, it is crucial to continue to hone and identify implementable and affordable solutions to detect true CDI. By using existing diagnostic pipelines and changing how tests are reported, change can be enacted much quicker with no need for additional monetary investment in equipment or physical infrastructure. The implementable components of this work could be combined with additional administrative strategies to improve CDI diagnosis. CDI is a clinical diagnosis and should be treated as such; clinicians need to always taking into consideration pre-test probability of disease prior to ordering diagnostic testing or initiating isolation or treatment [15]. At a system level, protocols implementation is necessary to prevent incorrectly ordered laboratory diagnostics to compliment any advances in understanding or more sophisticated

diagnostic strategies [16]. These kinds of multidisciplinary approaches could offer economical and highly effective approaches to improving diagnostics of CDI and other pathogens

The methods used in this work are not limited to CDI diagnostics and can be applied broadly to evaluate other opportunistic pathogens where qPCR diagnostics are used. If Ct probabilities were applied clinically for opportunistic pathogens, instead of simple dichotomous positive or negative results, there is a massive potential to decrease overdiagnosis/misdiagnosis leading to better antibiotic stewardship, lower healthcare costs, and most importantly improved patient outcomes and care.

4.5 References

1. Senchyna F, Gaur RL, Gombar S, Truong CY, Schroeder LF, Banaei N. *Clostridium difficile* PCR Cycle Threshold Predicts Free Toxin. Journal of Clinical Microbiology **2017**; 55(9): 2651-60.
2. Crobach MJT, Duszenko N, Terveer EM, Verduin CM, Kuijper EJ. Nucleic Acid Amplification Test Quantitation as Predictor of Toxin Presence in *Clostridium difficile* Infection. Journal of Clinical Microbiology **2018**; 56(3).
3. Kamboj M, Brite J, McMillen T, et al. Potential of real-time PCR threshold cycle (C-T) to predict presence of free toxin and clinically relevant *C. difficile* infection (CDI) in patients with cancer. Journal of Infection **2018**; 76(4): 369-75.
4. Davies KA, Planche T, Wilcox MH. The predictive value of quantitative nucleic acid amplification detection of *Clostridium difficile* toxin gene for faecal sample toxin status and patient outcome. Plos One **2018**; 13(12): 9.

5. Reigadas E, Alcala L, Valerio M, Marin M, Martin A, Bouza E. Toxin B PCR cycle threshold as a predictor of poor outcome of *Clostridium difficile* infection: a derivation and validation cohort study. Journal of Antimicrobial Chemotherapy **2016**; 71(5): 1380-5.
6. Origuen J, Orellana MA, Fernandez-Ruiz M, et al. Toxin B PCR Amplification Cycle Threshold Adds Little to Clinical Variables for Predicting Outcomes in *Clostridium difficile* Infection: a Retrospective Cohort Study. Journal of Clinical Microbiology **2019**; 57(2).
7. Rabe-Hesketh S, Skrondal A. Classical latent variable models for medical research. Statistical Methods in Medical Research **2008**; 17(1): 5-32.
8. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KGM, de Groot JAH. Latent Class Models in Diagnostic Studies When There is No Reference Standard-A Systematic Review. American Journal of Epidemiology **2014**; 179(4): 423-31.
9. Pepe MS, Janes H. Insights into latent class analysis of diagnostic test performance. Biostatistics **2007**; 8(2): 474-84.
10. Goetghebeur E, Liinev J, Boelaert M, Van der Stuyft P. Diagnostic test analyses in search of their gold standard: latent class analyses with random effects. Statistical Methods in Medical Research **2000**; 9(3): 231-48.
11. Rindskopf D, Rindskopf W. THE VALUE OF LATENT CLASS ANALYSIS IN MEDICAL DIAGNOSIS. Statistics in Medicine **1986**; 5(1): 21-7.
12. Polage CR, Gyorke CE, Kennedy MA, et al. Overdiagnosis of *Clostridium difficile* Infection in the Molecular Test Era. Jama Internal Medicine **2015**; 175(11): 1792-801.
13. Gabriel L, Beriot-Mathiot A. Hospitalization stay and costs attributable to *Clostridium difficile* infection: a critical review. Journal of Hospital Infection **2014**; 88(1): 12-21.

14. Pereira JA, McGeer A, Tomovici A, Selmani A, Chit A. Incidence and economic burden of *Clostridioides difficile* infection in Ontario: a retrospective population-based study. CMAJ open **2020**; 8(1): E16-E25.

15. Boly FJ, Reske KA, Kwon JH. The Role of Diagnostic Stewardship in *Clostridioides difficile* Testing: Challenges and Opportunities. Current Infectious Disease Reports **2020**; 22(4).

16. Liu C, Lan K, Krantz EM, et al. Improving Appropriate Diagnosis of *Clostridioides difficile* Infection Through an Enteric Pathogen Order Set With Computerized Clinical Decision Support: An Interrupted Time Series Analysis. Open Forum Infectious Diseases **2020**; 7(10).

17. Bosch DE, Mathias PC, Krumm N, Bryan A, Fang FC, Greninger AL. Elevated White Blood Cell Count Does Not Predict *Clostridium difficile* Nucleic Acid Testing Results. Clinical Infectious Diseases **2021**.

5.0 All References

1. Katz KC, Golding GR, Choi KB, et al. The evolving epidemiology of *Clostridium difficile* infection in Canadian hospitals during a post epidemic period (2009-2015). Canadian Medical Association Journal 2018; 190(25): E758-E65.

2. Bouza E. Consequences of *Clostridium difficile* infection: understanding the healthcare burden. Clinical Microbiology and Infection 2012; 18: 5-12.

3. Mylonakis E, Ryan ET, Calderwood SB. *Clostridium difficile*-associated diarrhea – A review. Archives of Internal Medicine 2001; 161(4): 525-33.

4. Guery B, Galperine T, Barbut F. *Clostridioides difficile*: diagnosis and treatments. Bmj-British Medical Journal 2019; 366: 19.

5. Daniels LM, Kufel WD. Clinical review of *Clostridium difficile* infection: an update on treatment and prevention. Expert Opinion on Pharmacotherapy 2018; 19(16): 1759-69.

6. Levy AR, Szabo SM, Lozano-Ortega G, et al. Incidence and Costs of *Clostridium difficile* Infections in Canada. Open Forum Infectious Diseases 2015; 2(3): 10.

7. Bouza E, Munoz P, Alonso R. Clinical manifestations, treatment and control of infections caused by *Clostridium difficile*. Clinical Microbiology and Infection 2005; 11: 57-64.

8. Postma N, Kiers D, Pickkers P. The challenge of *Clostridium difficile* infection: Overview of clinical manifestations, diagnostic tools and therapeutic options. International Journal of Antimicrobial Agents 2015; 46: S47-S50.

9. Korman TM. Diagnosis and Management of *Clostridium difficile* Infection. Seminars in Respiratory and Critical Care Medicine 2015; 36(1): 31-43.

10. Loo VaDIa. Association of Medical Microbiology and Infectious Disease Canada treatment practice guidelines for *Clostridium difficile* infection. Official Journal of the Association of Medical Microbiology and Infectious Disease Canada 2018; 3: 71-92.

11. McDonald LC, Gerding DN, Johnson S, et al. Clinical Practice Guidelines for *Clostridium difficile* Infection in Adults and Children: 2017 Update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). Clinical Infectious Diseases 2018; 66(7): 987-94.

12. Sandlund J, Wilcox MH. Ultrasensitive Detection of *Clostridium difficile* Toxins Reveals Suboptimal Accuracy of Toxin Gene Cycle Thresholds for Toxin Predictions. Journal of Clinical Microbiology 2019; 57(6): 5.

13. Garvey MI, Bradley CW, Wilkinson MAC, Holden E. Can a toxin gene NAAT be used to predict toxin EIA and the severity of *Clostridium difficile* infection? Antimicrobial Resistance and Infection Control 2017; 6: 8.

14. Senchyna F, Gaur RL, Gombar S, Truong CY, Schroeder LF, Banaei N. *Clostridium difficile* PCR Cycle Threshold Predicts Free Toxin. Journal of Clinical Microbiology 2017; 55(9): 2651-60.

15. Davies KA, Planche T, Wilcox MH. The predictive value of quantitative nucleic acid amplification detection of *Clostridium difficile* toxin gene for faecal sample toxin status and patient outcome. Plos One 2018; 13(12): 9.

16. Kamboj M, Brite J, McMillen T, et al. Potential of real-time PCR threshold cycle (C-T) to predict presence of free toxin and clinically relevant *C. difficile* infection (CDI) in patients with cancer. Journal of Infection 2018; 76(4): 369-75.

17. Dionne LL, Raymond F, Corbeil J, Longtin J, Gervais P, Longtin Y. Correlation between *Clostridium difficile* Bacterial Load, Commercial Real-Time PCR Cycle Thresholds, and Results of Diagnostic Tests Based on Enzyme Immunoassay and Cell Culture Cytotoxicity Assay. Journal of Clinical Microbiology 2013; 51(11): 3624-30.

18. Crobach MJT, Duszenko N, Terveer EM, Verduin CM, Kuijper EJ. Nucleic Acid Amplification Test Quantitation as Predictor of Toxin Presence in *Clostridium difficile* Infection. Journal of Clinical Microbiology 2018; 56(3).

19. Reigadas E, Alcala L, Valerio M, Marin M, Martin A, Bouza E. Toxin B PCR cycle threshold as a predictor of poor outcome of *Clostridium difficile* infection: a derivation and validation cohort study. Journal of Antimicrobial Chemotherapy 2016; 71(5): 1380-5.

20. Origuen J, Orellana MA, Fernandez-Ruiz M, et al. Toxin B PCR Amplification Cycle Threshold Adds Little to Clinical Variables for Predicting Outcomes in *Clostridium difficile* Infection: A Retrospective Cohort Study. Journal of Clinical Microbiology 2019; 57(2).

21. Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: the case of tests with continuous results. Biochemia Medica 2016; 26(3): 297-307.

22. Habibzadeh F, Habibzadeh P. The likelihood ratio and its graphical representation. Biochemia Medica 2019; 29(2): 7.

23. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KGM, de Groot JAH. Latent Class Models in Diagnostic Studies When There is No Reference Standard-A Systematic Review. American Journal of Epidemiology 2014; 179(4): 423-31.

24. Visser I, Speekenbrink M. depmixS4: An R Package for Hidden Markov Models. Journal of Statistical Software 2010; 36(7): 1-21.

1. Centers for Disease Control and Prevention. Antibiotic resistance threats in the United States. Available at: https://www.cdc.gov/drugresistance/biggest-threats.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fdrugresistance%2Fbiggest_threats.html#cdiff. Accessed 27 October 2020.
2. Peng Z, Ling LF, Stratton CW, et al. Advances in the diagnosis and treatment of Clostridium difficile infections. Emerg Microbes Infect 2018; 7:15.
3. Daniels LM, Kufel WD. Clinical review of *Clostridium difficile* infection: an update on treatment and prevention. Expert Opin Pharmacother 2018; 19:1759–69.
4. Crobach MJT, Vernon JJ, Loo VG, et al. Understanding *Clostridium difficile* colonization. Clin Microbiol Rev 2018; 31:29.
5. Burnham CA, Carroll KC. Diagnosis of *Clostridium difficile* infection: an ongoing

conundrum for clinicians and for clinical laboratories. Clin Microbiol Rev 2013; 26:604–30.

6. Bartsch SM, Umscheid CA, Nachamkin I, Hamilton K, Lee BY. Comparing the economic and health benefits of different approaches to diagnosing *Clostridium difficile* infection. Clin Microbiol Infect 2015; 21:77.e1–9.

7. McDonald LC, Gerding DN, Johnson S, et al. Clinical practice guidelines for Clostridium difficile infection in adults and children: 2017 update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). Clin Infect Dis 2018; 66:987–94.

8. Guery B, Galperine T, Barbut F. *Clostridioides difficile*: diagnosis and treatments. BMJ 2019; 366:19.

9. Hanania A, Jiang ZD, Smiley C, Lasco T, Garey KW, DuPont HL. Fecal calprotectin in the diagnosis of *Clostridium difficile* infection. Infect Dis Clin Pract 2016; 24:31–4.

10. Popiel KY, Gheorghe R, Eastmond J, Miller MA. Usefulness of adjunctive fecal calprotectin and serum procalcitonin in individuals positive for *Clostridium difficile* toxin gene by PCR assay. J Clin Microbiol 2015; 53:3667–9.

11. Kim HN, Kim H, Moon HW, Hur M, Yun YM. Toxin positivity and tcdB gene load in broad-spectrum *Clostridium difficile* infection. Infection 2018; 46:113–7.

12. Polage CR, Gyorke CE, Kennedy MA, et al. Overdiagnosis of *Clostridium difficile* infection in the molecular test era. JAMA Intern Med 2015; 175:1792–801.

13. Hitchcock MM, Holubar M, Hogan CA, Tompkins LS, Banaei N. Dual reporting of *Clostridioides difficile* PCR and predicted toxin result based on PCR cycle

threshold reduces treatment of toxin-negative patients without increases in adverse outcomes. J Clin Microbiol 2019; 57:11.

14. Kamboj M, Brite J, McMillen T, et al. Potential of real-time PCR threshold cycle (C-T) to predict presence of free toxin and clinically relevant *C. difficile* infection (CDI) in patients with cancer. J Infect 2018; 76:369–75.

15. Dionne LL, Raymond F, Corbeil J, Longtin J, Gervais P, Longtin Y. Correlation between *Clostridium difficile* bacterial load, commercial real-time PCR cycle thresholds, and results of diagnostic tests based on enzyme immunoassay and cell culture cytotoxicity assay. J Clin Microbiol 2013; 51:3624–30.

16. Crobach MJT, Duszenko N, Terveer EM, Verduin CM, Kuijper EJ. Nucleic acid amplification test quantitation as predictor of toxin presence in *Clostridium difficile* infection. J Clin Microbiol 2018; 56:e01316–17.

17. Davies KA, Planche T, Wilcox MH. The predictive value of quantitative nucleic acid amplification detection of *Clostridium difficile* toxin gene for faecal sample toxin status and patient outcome. PloS One 2018; 13:9.

18. Garvey MI, Bradley CW, Wilkinson MAC, Holden E. Can a toxin gene NAAT be used to predict toxin EIA and the severity of *Clostridium difficile* infection? Antimicrob Resist Infect Control 2017; 6:8.

19. Senchyna F, Gaur RL, Gombar S, Truong CY, Schroeder LF, Banaei N. *Clostridium difficile* PCR cycle threshold predicts free toxin. J Clin Microbiol 2017; 55:2651–60.

20. Jazmati N, Hellmich M, Ličanin B, Plum G, Kaasch AJ. PCR cycle threshold value predicts the course of *Clostridium difficile* infection. Clin Microbiol Infect 2016;

22:e7–8.

21. Reigadas E, Alcalá L, Valerio M, Marín M, Martin A, Bouza E. Toxin B PCR cycle threshold as a predictor of poor outcome of *Clostridium difficile* infection: a derivation and validation cohort study. J Antimicrob Chemother 2016; 71:1380–5.

22. Halligan S, Altman DG, Mallett S. Disadvantages of using the area under the receiver operating characteristic curve to assess imaging tests: a discussion and proposal for an alternative approach. Eur Radiol 2015; 25:932–9.

23. Origuen J, Orellana MA, Fernandez-Ruiz M, et al. Toxin B PCR amplification cycle threshold adds little to clinical variables for predicting outcomes in *Clostridium difficile* infection: a retrospective cohort study. J Clin Microbiol 2019; 57:e01125–18.

24. Sandlund J, Wilcox MH. Ultrasensitive detection of *Clostridium difficile* toxins reveals suboptimal accuracy of toxin gene cycle thresholds for toxin predictions. J Clin Microbiol 2019; 57:5.

25. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KG, de Groot JA. Latent class models in diagnostic studies when there is no reference standard—a systematic review. Am J Epidemiol 2014; 179:423–31.

26. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. Am J Epidemiol 1995; 141:263–72.

27. Weakliem DL. A critique of the Bayesian information criterion for model selection. Sociol Methods Res 1999; 27:359–97.

28. Visser I, Speekenbrink M. depmixS4: an R package for Hidden Markov models. J

Stat Softw 2010; 36:1–21.

29. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2019.

30. Mweu MM, Wambua J, Njuga F, Bejon P, Mwanga D. Bayesian evaluation of the performance of three diagnostic tests for Plasmodium falciparum infection in a low-transmission setting in Kilifi County, Kenya. Wellcome Open Res 2019; 4:67.

31. Stout JE, Wu Y, Ho CS, et al; Tuberculosis Epidemiologic Studies Consortium. Evaluating latent tuberculosis infection diagnostics using latent class analysis. Thorax 2018; 73:1062–70.

32. Clements MN, Corstjens P, Binder S, et al. Latent class analysis to evaluate performance of point-of-care CCA for low-intensity Schistosoma mansoni infections in Burundi. Parasit Vectors 2018; 11:13.

33. Jespers V, Menten J, Smet H, et al. Quantification of bacterial species of the vaginal microbiome in different groups of women, using nucleic acid amplification tests. BMC Microbiol 2012; 12:10.

34. Sigera PC, Amarasekara R, Rodrigo C, et al. Risk prediction for severe disease and better diagnostic accuracy in early dengue infection; the Colombo dengue study. BMC Infect Dis 2019; 19:8.

35. Castro-Sesquen YE, Saldana A, Patino Nava D, et al. Use of a latent class analysis in the diagnosis of chronic Chagas disease in the Washington metropolitan area. Clin Infect Dis 2020. Doi:10.1093/cid/ciaa1101

36. Baughman AL, Bisgard KM, Cortese MM, Thompson WW, Sanden GN,

Strebel PM. Utility of composite reference standards and latent class analysis in evaluating the clinical accuracy of diagnostic tests for pertussis. Clin Vaccine Immunol 2008; 15:106–14.

1. Magill SS, Edwards JR, Bamberg W, et al. Multistate Point- Prevalence Survey of Health Care- Associated Infections. New England Journal of Medicine **2014**; 370(13): 1198-208.
2. Crobach MJT, Planche T, Eckert C, et al. European Society of Clinical Microbiology and Infectious Diseases: update of the diagnostic guidance document for Clostridium difficile infection. Clinical Microbiology and Infection **2016**; 22: S63-S81.
3. Daniels LM, Kufel WD. Clinical review of *Clostridium difficile* infection: an update on treatment and prevention. Expert Opinion on Pharmacotherapy **2018**; 19(16): 1759-69.
4. Bartsch SM, Umscheid CA, Nachamkin I, Hamilton K, Lee BY. Comparing the economic and health benefits of different approaches to diagnosing *Clostridium difficile* infection. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases **2015**; 21(1): 77.e1-9.
5. Crobach MJT, Baktash A, Duszenko N, Kuijper EJ. Diagnostic Guidance for *C. difficile* Infections. In: Mastrantonio P, Rupnik M. Updates on Clostridium Difficile in Europe: Advances in Microbiology, Infectious Diseases and Public Health, Vol 8 Vol. 1050, **2018**:27-44.
6. Planche TD, Davies KA, Coen PG, et al. Differences in outcome according to Clostridium difficile testing method: a prospective multicentre diagnostic validation study of *C difficile* infection. Lancet Infectious Diseases **2013**; 13(11): 936-45.

7. Guery B, Galperine T, Barbut F. *Clostridioides difficile*: diagnosis and treatments. Bmj-British Medical Journal **2019**; 366: 19.

8. Rao KN, Malani PN. Diagnosis and Treatment of *Clostridioides (Clostridium) difficile* Infection in Adults in 2020. Jama-Journal of the American Medical Association **2020**; 323(14): 1403-4.

9. Kamboj M, Brite J, McMillen T, et al. Potential of real-time PCR threshold cycle (C-T) to predict presence of free toxin and clinically relevant *C. difficile* infection (CDI) in patients with cancer. Journal of Infection **2018**; 76(4): 369-75.

10. Dionne LL, Raymond F, Corbeil J, Longtin J, Gervais P, Longtin Y. Correlation between *Clostridium difficile* Bacterial Load, Commercial Real-Time PCR Cycle Thresholds, and Results of Diagnostic Tests Based on Enzyme Immunoassay and Cell Culture Cytotoxicity Assay. Journal of Clinical Microbiology **2013**; 51(11): 3624-30.

11. Kim HN, Kim H, Moon HW, Hur M, Yun YM. Toxin positivity and tcdB gene load in broad-spectrum *Clostridium difficile* infection. Infection **2018**; 46(1): 113-7.

12. Crobach MJT, Duszenko N, Terveer EM, Verduin CM, Kuijper EJ. Nucleic Acid Amplification Test Quantitation as Predictor of Toxin Presence in *Clostridium difficile* Infection. Journal of Clinical Microbiology **2018**; 56(3).

13. Senchyna F, Gaur RL, Gombar S, Truong CY, Schroeder LF, Banaei N. *Clostridium difficile* PCR Cycle Threshold Predicts Free Toxin. Journal of Clinical Microbiology **2017**; 55(9): 2651-60.

14. Davies KA, Planche T, Wilcox MH. The predictive value of quantitative nucleic acid amplification detection of *Clostridium difficile* toxin gene for faecal sample toxin status and patient outcome. Plos One **2018**; 13(12): 9.

15. Garvey MI, Bradley CW, Wilkinson MAC, Holden E. Can a toxin gene NAAT be used to predict toxin EIA and the severity of *Clostridium difficile* infection? Antimicrobial Resistance and Infection Control **2017**; 6: 8.

16. Jazmati N, Hellmich M, Licanin B, Plum G, Kaasch AJ. PCR cycle threshold value predicts the course of *Clostridium difficile* infection. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases **2016**; 22(2): e7-e8.

17. Reigadas E, Alcala L, Valerio M, Marin M, Martin A, Bouza E. Toxin B PCR cycle threshold as a predictor of poor outcome of *Clostridium difficile* infection: a derivation and validation cohort study. Journal of Antimicrobial Chemotherapy **2016**; 71(5): 1380-5.

18. Doolan CP, Louie T, Lata C, et al. Latent class analysis for the diagnosis of *Clostridioides difficile* infection. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **2020**.

19. Joseph L, Gyorkos TW, Coupal L. BAYESIAN-ESTIMATION OF DISEASE PREVALENCE AND THE PARAMETERS OF DIAGNOSTIC-TESTS IN THE ABSENCE OF A GOLD STANDARD. American Journal of Epidemiology **1995**; 141(3): 263-72.

20. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KGM, de Groot JAH. Latent Class Models in Diagnostic Studies When There is No Reference Standard-A Systematic Review. American Journal of Epidemiology **2014**; 179(4): 423-31.

21. Humphries RM, Uslan DZ, Rubin Z. Performance of *Clostridium difficile* Toxin Enzyme Immunoassay and Nucleic Acid Amplification Tests Stratified by Patient Disease Severity. Journal of Clinical Microbiology **2013**; 51(3): 869-73.

22. McDonald LC, Gerding DN, Johnson S, et al. Clinical Practice Guidelines for *Clostridium difficile* Infection in Adults and Children: 2017 Update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). Clinical Infectious Diseases **2018**; 66(7): E1-E48.

23. Abou Chakra CN, Pepin J, Sirard S, Valiquette L. Risk Factors for Recurrence, Complications and Mortality in *Clostridium difficile* Infection: A Systematic Review. Plos One **2014**; 9(6): 14.

24. di Masi A, Leboffe L, Polticelli F, et al. Human Serum Albumin Is an Essential Component of the Host Defense Mechanism Against *Clostridium difficile* Intoxication. Journal of Infectious Diseases **2018**; 218(9): 1424-35.

25. Hardt C, Berns T, Treder W, Dumoulin FL. Univariate and multivariate analysis of risk factors for severe *clostridium difficile*-associated diarrhoea: Importance of co-morbidity and serum C-reactive protein. World Journal of Gastroenterology **2008**; 14(27): 4338-41.

26. See I, Mu Y, Cohen J, et al. NAP1 Strain Type Predicts Outcomes From *Clostridium difficile* Infection. Clinical Infectious Diseases **2014**; 58(10): 1394-400.

27. Team RC. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, **2019**.

28. Origuen J, Orellana MA, Fernandez-Ruiz M, et al. Toxin B PCR Amplification Cycle Threshold Adds Little to Clinical Variables for Predicting Outcomes in *Clostridium difficile* Infection: a Retrospective Cohort Study. Journal of Clinical Microbiology **2019**; 57(2).

29. Bosch DE, Mathias PC, Krumm N, Bryan A, Fang FC, Greninger AL. Elevated White Blood Cell Count Does Not Predict *Clostridium difficile* Nucleic Acid Testing Results. Clinical Infectious Diseases **2021**.

30. Pollock NR, Banz A, Chen XH, et al. Comparison of *Clostridioides difficile* Stool Toxin Concentrations in Adults With Symptomatic Infection and Asymptomatic Carriage Using an Ultrasensitive Quantitative Immunoassay. Clinical Infectious Diseases **2019**; 68(1): 78-86.

31. El Feghaly RE, Stauber JL, Deych E, Gonzalez C, Tarr PI, Haslam DB. Markers of Intestinal Inflammation, Not Bacterial Burden, Correlate With Clinical Outcomes in *Clostridium difficile* Infection. Clinical Infectious Diseases **2013**; 56(12): 1713-21.

32. Kelly CP, Chen XH, Williams D, et al. Host Immune Markers Distinguish *Clostridioides difficile* Infection From Asymptomatic Carriage and Non-C. difficile Diarrhea. Clinical Infectious Diseases **2020**; 70(6): 1083-93.

1. Senchyna F, Gaur RL, Gombar S, Truong CY, Schroeder LF, Banaei N. *Clostridium difficile* PCR Cycle Threshold Predicts Free Toxin. Journal of Clinical Microbiology **2017**; 55(9): 2651-60.

2. Crobach MJT, Duszenko N, Terveer EM, Verduin CM, Kuijper EJ. Nucleic Acid Amplification Test Quantitation as Predictor of Toxin Presence in *Clostridium difficile* Infection. Journal of Clinical Microbiology **2018**; 56(3).

3. Kamboj M, Brite J, McMillen T, et al. Potential of real-time PCR threshold cycle (C-T) to predict presence of free toxin and clinically relevant *C. difficile* infection (CDI) in patients with cancer. Journal of Infection **2018**; 76(4): 369-75.

4. Davies KA, Planche T, Wilcox MH. The predictive value of quantitative nucleic acid amplification detection of *Clostridium difficile* toxin gene for faecal sample toxin status and patient outcome. Plos One **2018**; 13(12): 9.

5. Reigadas E, Alcala L, Valerio M, Marin M, Martin A, Bouza E. Toxin B PCR cycle threshold as a predictor of poor outcome of *Clostridium difficile* infection: a derivation and validation cohort study. Journal of Antimicrobial Chemotherapy **2016**; 71(5): 1380-5.

6. Origuen J, Orellana MA, Fernandez-Ruiz M, et al. Toxin B PCR Amplification Cycle Threshold Adds Little to Clinical Variables for Predicting Outcomes in *Clostridium difficile* Infection: a Retrospective Cohort Study. Journal of Clinical Microbiology **2019**; 57(2).

7. Rabe-Hesketh S, Skrondal A. Classical latent variable models for medical research. Statistical Methods in Medical Research **2008**; 17(1): 5-32.

8. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KGM, de Groot JAH. Latent Class Models in Diagnostic Studies When There is No Reference Standard-A Systematic Review. American Journal of Epidemiology **2014**; 179(4): 423-31.

9. Pepe MS, Janes H. Insights into latent class analysis of diagnostic test performance. Biostatistics **2007**; 8(2): 474-84.

10. Goetghebeur E, Liinev J, Boelaert M, Van der Stuyft P. Diagnostic test analyses in search of their gold standard: latent class analyses with random effects. Statistical Methods in Medical Research **2000**; 9(3): 231-48.

11. Rindskopf D, Rindskopf W. THE VALUE OF LATENT CLASS ANALYSIS IN MEDICAL DIAGNOSIS. Statistics in Medicine **1986**; 5(1): 21-7.

12. Polage CR, Gyorke CE, Kennedy MA, et al. Overdiagnosis of *Clostridium difficile* Infection in the Molecular Test Era. Jama Internal Medicine **2015**; 175(11): 1792-801.

13. Gabriel L, Beriot-Mathiot A. Hospitalization stay and costs attributable to *Clostridium difficile* infection: a critical review. Journal of Hospital Infection **2014**; 88(1): 12-21.

14. Pereira JA, McGeer A, Tomovici A, Selmani A, Chit A. Incidence and economic burden of *Clostridioides difficile* infection in Ontario: a retrospective population-based study. CMAJ open **2020**; 8(1): E16-E25.

15. Boly FJ, Reske KA, Kwon JH. The Role of Diagnostic Stewardship in *Clostridioides difficile* Testing: Challenges and Opportunities. Current Infectious Disease Reports **2020**; 22(4).

16. Liu C, Lan K, Krantz EM, et al. Improving Appropriate Diagnosis of *Clostridioides difficile* Infection Through an Enteric Pathogen Order Set With Computerized Clinical Decision Support: An Interrupted Time Series Analysis. Open Forum Infectious Diseases **2020**; 7(10).

6.0 Appendix

6.1 Author's statement for letter of permission:

Signatures removed

I allow Cody P. Doolan to use the materials of the paper titled "Latent Class Analysis for the Diagnosis of *Clostridioides difficile* Infection" for his graduate studies thesis.

_____

Dylan Pillai

_____

Thomas Louie

_____

Christopher Lata

_____

Oscar Larios

_____

William Stokes

_____

Joseph Kim

_____

Kristen Brown

_____

Paul Beck Dated

_____

Rob Deardon Dated