Mathematical Models in Medicine: The Immune Response of Celiac Disease and the Environmental Transmission of *Clostridioides difficile* in Healthcare Settings

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Cara Jill Sulyok

May 2021

For my sister, Leah.

Acknowledgments

I could not have completed this work nor written these pages without the countless individuals who have inspired and encouraged me. Words cannot express the gratitude that I have for you all being integral in my life.

Thank you to my two phenomenal advisors, Suzanne Lenhart and Judy Day, for their guidance, encouragement, and dedication not only in this research, but throughout my graduate school career. From the beginning when I shared that I wanted to mathematically model celiac disease, you both have been incredibly supportive. Thank you being my role models; I am humbled and honored to work with both of you and hope to inspire other students as you have inspired me.

Thank you to the members of my committee, Steven Wise and Louis Gross, for your time, thoughtful feedback, and willingness in evaluating my work, not only with this dissertation, but in the classroom.

I am very appreciative to my collaborators Lindsey Fox and Cristina Lanzas for their work, opinions, and advice in our *Clostridioides difficile* project and helping to prepare our results for publication.

To Joseph Murray, thank you for sharing your knowledge on the immune response in celiac disease and providing insight on the Q-MARSH diagnostic method.

To Malissa Meadows, thank you for recognizing my talents, serving as my teaching mentor, and allowing me to design courses with you. I've grown so much as an educator during my time at UTK and am incredibly grateful that I was able to work closely with you.

To Pam Armentrout, thank you for your open door and warm hugs. I always felt like I belonged in our department because of you.

To Mo Yahdi, thank you for believing in me, igniting my love of interdisciplinary research, and helping me navigate my transition to graduate school.

To Sally Lister, thank you for being the first to recognize my talent in mathematics and encouraging me to pursue mathematics as my future.

And last, but certainly not least, thanks needs to be extended to my family and friends. To my mom and dad, thank you for developing my passion for learning and teaching me that I can achieve anything I put my mind to. More than that, thank you for your love, never-ending support, and sincere belief in me. You have motivated me in my academic journey and life to achieve my goals and chase my dreams. To my sister, Leah, thank you for the constant support and friendship. You inspire me to be better in all aspects of my life. Also, your Tika and Pyewacket pictures got me through the many difficult days of graduate school. To my friends and fellow graduate students, there are far too many of you to name, but you know who you are. Thank you for the conversations, collaborations, and distractions; I'll miss being down the hall from all of you. To Tricia, my graduate school day one, thank you for the laughs, the tears, and the adventures. I wouldn't have made it through our first year or the years following without you. And to Chris, thank you for the hikes when I needed the fresh air, for the Indian food when I needed nourishment, for being there when I needed a hug, and everything in between - I am glad that you're in my life.

Thank you all for believing in me and helping me to accomplish this milestone, especially on the days that I doubted myself! I'm excited for what the future holds and know that you will all continue to be my support system; I'm lucky to have all of you.

The work related to *Clostridioides difficile* in Chapter 3 was supported by the joint Division of Mathematical Sciences and National Institute of General Medical Sciences Mathematical Biology Program through the National Institute of Health award #R01GM113239.

It doesn't matter what it is. What matters is what it will become.

- Dr. Seuss

Abstract

Mathematical modeling is a useful technique to describe dynamics happening within events and allows one to address questions and test hypotheses that may be not be feasible to study in reality. This work uses mathematical models to describe two such phenomena, one relating to immunology and the other to the spread of infectious diseases.

Celiac disease is a hereditary autoimmune disease that affects approximately 1 in 133 Americans. It is caused by a reaction to the protein gluten found in wheat, rye, and barley. After ingesting gluten, a patient with celiac disease may experience a range of unpleasant symptoms while small intestinal villi, essential to nutrient absorption, are destroyed in an immune-mediated process. The only known treatment for this disease is a lifelong gluten-free diet and there is currently no drug treatment. The first study provides a mathematical framework to better understand the effects of immune activation on gut health. This mathematical model uses a system of ordinary differential equations to track changes in small intestinal cell densities and relates them to the Q-MARSH score, a criterion used in the diagnosis of celiac disease. The model can be used to investigate and analyze the immune response and various theories behind the progression of this disease by focusing on understanding the dynamics of the small intestine in situations mirroring healthy function and celiac disease. By doing so, we can assist in further quantifying and augmenting diagnostic measures and investigate potential therapies to mitigate the negative effects of celiac disease.

Clostridioides difficile (*C. difficile*) is the leading cause of infectious diarrhea and one of the most frequently identified healthcare-acquired infections in United States hospitals. *C. difficile* is typically contracted after antibiotic use, when healthy gut microbiota that prevent colonization is compromised. Colonized patients, both symptomatic and

asymptomatic, shed *C. difficile* endospores that can survive for long periods on surfaces outside the host and are resistant to many commonly-used disinfectants. Transmission pathways can include contact with endospores on fomites, objects likely to carry infection. The second modeling study investigates the relative contribution of two environmental pathways to *C. difficile* transmission in healthcare settings. Due to the small hospital ward size, patient and pathogen populations are simulated stochastically and compared with the average behavior described by a system of ordinary differential equations. The results can be utilized to examine the role surfaces with varying touch frequencies contribute to patient colonization with *C. difficile* in healthcare settings.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Mathematical modeling allows one to describe dynamics happening in the real world by translating the knowledge and beliefs of interactions into the language of mathematics. For this reason, mathematical modeling is a useful technique used to describe natural occurrences and allows one to address questions and test hypotheses that may be not be feasible to study in reality. Modeling interacting systems of the real world requires compromise. In order to avoid complicated systems, mathematical modelers must first identify the most important components in the system to include in the model, excluding everything else. Even with these simplifying assumptions, mathematical models can be used to develop scientific understanding, test the effect of changes in a system, and aid in decision making. This work uses mathematical models to better understand dynamics in two different systems, one relating to immunology and the other to the spread of infectious diseases.

Celiac disease is a hereditary autoimmune disease that affects approximately 1 in 133 Americans. It is caused by a reaction to the protein gluten found in wheat, rye, and barley. After ingesting gluten, a patient with celiac disease may experience a range of unpleasant symptoms while small intestinal villi, essential to nutrient absorption, are destroyed in an immune-mediated process. The only known treatment for this disease is a lifelong gluten-free diet and there is currently no drug treatment.

The first study provides a mathematical framework to better understand the effects of immune activation on gut health. This mathematical model uses a four-dimensional system of ordinary differential equations (ODEs) to track changes in small intestinal cell densities

and relates them to the Q-MARSH score, a criterion used in the diagnosis of celiac disease. The four compartments include the cells in both the base and tip of the villus and the cells in the crypt representing the intestinal structure, as well as the intraepithelial lymphocytes (IELs) representing the immune response. In Chapter 2, we introduce our mathematical model and simulate individuals with varying levels of gut health to understand both short- and long-term behavior of patients suffering from celiac disease. To do so, we include multiple profiles of the immune response through varying activation functions for the IELs, including early, delayed, and damped activation, as well as both relatively weak and strong levels of response. Additionally, patient-level data from gluten challenges is used to determine unknown parameter rates involved in the immune response. Knowledge about the drivers of small intestinal cellular turnover can be used to investigate and analyze the immune response as well as better understand various theories behind the progression of this disease. The model is able to replicate the dynamics of the small intestine in situations mirroring healthy function and celiac disease. By doing so, we can assist in further quantifying and augmenting diagnostic measures and investigate potential therapies to mitigate the negative effects of celiac disease.

Clostridioides difficile (*C. difficile*) is the leading cause of infectious diarrhea and the most frequently identified healthcare-associated infection in United States hospitals [41]. *C. difficile* is typically contracted after antibiotic use, when healthy gut microbiota that prevent colonization is compromised. Colonized patients, both symptomatic and asymptomatic, shed *C. difficile* endospores that can survive for long periods on surfaces outside the host and are resistant to many commonly-used disinfectants [58]. These spores can be transmitted to susceptible patients through contact with contaminated hands and fomites, objects or surfaces that can harbor infectious agents. Fomites can be classified as high-touch or low-touch based on the frequency they are contacted. The second modeling study investigates the relative contribution of these two environmental pathways on new cases of *C. difficile* colonization among patients of a hospital ward. In Chapter 3, we extend a published model of *C. difficile* to include high-touch and low-touch frequency surfaces to quantify their roles in the transmission of *C. difficile* in a hospital ward. High-touch surfaces, which are interacted with more often and therefore receive extra cleaning, include

bed rails, tray tables, IV poles, doorknobs, sinks, and toilets. Conversely, low-touch surfaces are handled less often and thus cleaned less frequently; some examples include curtains, furniture, thermostats, soap and paper towel dispensers, and trash cans [31]. The dynamics of transmission are described by a system of ODEs representing patient population classes and pathogen environmental reservoirs. There are four patient compartments, including resistant, susceptible, colonized (asymptomatic), and diseased (symptomatic) individuals, as well as two environmental compartments, represented by high-touch and low-touch frequency surfaces.

Due to the small hospital ward size, patient and pathogen populations are simulated stochastically and compared with the average behavior described by the system of ODEs. Deterministic ODE models generate the average behavior of the system. However, a particular outbreak of *C. difficile* in a hospital ward may not necessarily follow the behavior of an average outbreak. Thus, the events modeled by the ODE were simulated using the Gillespie Stochastic Simulation Algorithm (GSSA) [27] to understand the variability among different trajectories, or time courses, of the system. The GSSA views interactions between populations as discrete events that occur randomly. Based only on the current population levels, not past levels, the time step and the event that occurs after that time step are determined probabilistically by propensity functions. Population levels are then updated by integer values based on the event that occurred. As such, the probabilistic occurrence of interactions among the six state variables are involved in 15 different GSSA events.

A global sensitivity analysis conducted on the ODE model helped to determine which parameters have a significant impact on incidence. In particular, there were three outcomes of interest, including

1. the incidence of colonized patients due to contact with high-touch fomites,
2. the incidence of colonized patients due to contact with low-touch fomites, and
3. the incidence of diseased patients.

Parameter values were sampled using Latin Hypercube Sampling, Partial Rank Correlation Coefficients (PRCC) were computed as the sensitivity indices, and associated p-values were

calculated to determine the significance of each PRCC value. The PRCC values and p-values for each parameter relative to the three outputs of interest aid in quantifying each parameter's contribution to each outcome and identifying feasible strategies to prevent disease transmission.

Several scenarios considering changes in the values of these impactful parameters were stochastically simulated to determine the factors that increase incidence. The results show that increased contacts with high-touch fomites cause the largest increase in the incidence of colonizations while an increased rate at which colonized patients become diseased is responsible for the largest increase in the incidence of diseased. It is important to note that from the formulation of the model, only the incidence of colonized patients can be linked with the environmental reservoir class that lead to colonization. The results can be utilized to examine the role surfaces with varying touch frequencies contribute to patient colonization with *C. difficile* in healthcare settings.

Chapter 2

A Mathematical Framework to Augment the Q-MARSH Score in the Diagnosis of Celiac Disease

2.1 Introduction

The overall function of the immune system is to prevent or limit infection. The immune system is the body's defense against infectious organisms and other invaders. Through a process called the immune response, the immune system attacks organisms and substances that invade the body and cause disease.

The immune system has the ability to distinguish between normal, healthy cells and unhealthy cells by recognizing a variety of "danger" cues called danger-associated molecular patterns. Cells may be unhealthy due to infection or because of cellular damage caused by non-infectious agents such as sunburn or cancer. Infectious microbes such as viruses and bacteria release another set of signals recognized by the immune system called pathogen-associated molecular patterns. When the immune system recognizes these signals, it responds to address the perceived problem. If an immune response cannot be activated when there is sufficient need, complications arise, including infection. Conversely, when an immune

response is activated without a real threat or is not turned off once the problem subsides, different complications arise, including allergic reactions and autoimmune diseases [59].

The immune system is complex and extensive. There are numerous cell types that either circulate throughout the body or reside in a particular tissue. Each cell type plays a unique role, with different methods of recognizing problems, communicating with other cells, and performing their functions [59]. By understanding all the details behind this network, researchers may optimize immune responses to confront specific issues, ranging from infections to cancer.

2.2 Autoimmune Disorders

Disorders of the immune system fall into four main categories - (1) immunodeficiency disorders, either primary or acquired, (2) autoimmune disorders, in which the body's immune system attacks its own tissue as a foreign matter, (3) allergic disorders, in which the immune system overreacts in response to an antigen, and (4) cancers of the immune system [59]. Our focus will be on a particular autoimmune disorder involved with food ingestion, celiac disease, and related diseases collectively referred to as "leaky gut."

Abnormal reactions following food ingestion are commonly defined as "adverse reactions to food," representing a variety of immune-mediated food allergies and non-immune-mediated food intolerances. Food allergy and food hypersensitvitiy are terms used interchangeably, but are reserved for those reactions mediated by the immune system. Conversely, most food intolerances are caused by enzyme deficiencies, as in case of lactose intolerance, without the direct involvement of the immune system [59].

In autoimmune disorders, the immune system mistakenly attacks the body's healthy organs and tissues as though they were foreign invaders. Common autoimmune disorders include Crohn's disease, celiac disease, Graves' disease, Hashimoto's thyroiditis, systemic lupus erythematosus, multiple sclerosis, rheumatoid arthritis, and type 1 diabetes. Autoimmune diseases can range from targeting only one organ, as is the case of type 1 diabetes damaging the pancreas; whereas other diseases, like systemic lupus erythematosus, effect the whole body.

Autoimmune disorders differ from allergies in the type of immune response. In an autoimmune response, tissue destruction occurs. With allergies, the immune system overreacts to harmless allergens [59].

Celiac disease (CeD) is an acquired chronic autoimmune disorder that develops in genetically predisposed individuals in which the patient lacks the ability to tolerate ingested prolamin glycoproteins found in wheat (gliadins) and related derivatives of barley (hordeins) and rye (secalins), collectively referred to as "gluten" [1]. CeD is also known as celiac sprue, non-tropical sprue, idiopathic sprue, idiopathic steatorrhoea, and gluten-sensitive enteropathy [22]. Innate and adaptive immunologic reactions to gluten result in chronic inflammatory responses in the small intestinal mucosa resulting in both structural and functional abnormalities [1]. CeD affects approximately 1 in 133 Americans with a similar incidence worldwide, though this ratio is speculated to be much higher.

Those who suffer from CeD, often referred to as celiacs, suffer the damaging effects of ingested gluten to the intestinal villi, the long, finger-like structures which protrude from the intestinal wall in the small intestine. This histologic damage negatively impacts the resistance of the small intestine. As the resistance of the small intestine declines, gluten may pass through the wall of the small intestine inducing an immune response which further contributes to the degeneration of the small intestine. Upon entry of gluten into the small intestine, a patient develops painful digestion disorders due to villi impairment and loss of absorption, leading to ailments such as vitamin deficiencies and osteoporosis [21].

2.3 Background on Celiac Disease

This work provides a mathematical framework to better understand the biological and immunological mechanisms in celiac disease. We formulate a mathematical model with immunologic and intestinal features using a system of ordinary differential equations (ODEs) and relate these dynamics to typical diagnostic histologic measurements. The model will be able to analyze various theories behind the progression of this disease by mirroring the cellular dynamics of a healthy subject and a patient suffering from CeD. By doing so, we can suggest potential therapies to mitigate the effects and potential progression of the disease.

Much of this research is limited in a laboratory setting as progression of the disease to a state where a gluten-free diet (GFD) no longer moderates the immune response can only be observed with invasive procedures. A mathematical model can make predictions and guide laboratory experiments to determine potential and practical therapies.

2.3.1 Symptoms

The symptoms of CeD are widely varied as some patients have no evident symptoms while others experience severe symptoms. When symptoms are present, they generally include some combination of gastrointestinal (e.g., diarrhea and abdominal discomfort) and system manifestations (e.g., dermatitis herpetiformis). Nutritional deficiencies (e.g., iron deficiency) may also lead to symptoms such as tiredness and fatigue [1]. The most common symptoms are abdominal pain, diarrhea, weight loss, anemia, and bone or joint pain. Some less common symptoms include fatigue, depression, anxiety, seizures, and infertility [11]. Undiagnosed, these symptoms may worsen over time, so early diagnosis is imperative. Unfortunately, these symptoms are common to other ailments, making proper diagnosis difficult. Additionally, some of those suffering from CeD exhibit no symptoms, further complicating detection. Asymptomatic or minimally symptomatic patients can still exhibit severe histologic structural changes [1].

2.3.2 Treatment

Currently, the only known and accepted treatment for CeD is a strict, lifelong compliance to a GFD, which consists of the removal of all gluten-containing products, such as breads, cakes, cereals, cookies, crackers, pastas, and soups. There is no current drug treatment for CeD. For most patients with CeD, avoiding gluten will alleviate their symptoms and allow the small intestine to repair. Many foods are naturally gluten-free, including fresh fruits and vegetables, rice, dairy products, and organic meats. However, maintaining a GFD can be demanding. While the United States Food and Drug Administration (FDA) requires that wheat be identified on packaging, many products, especially health and beauty products, do not identify gluten as a constituent. “Gluten-free” is a voluntary claim that can be used by

food manufacturers on food labels if they meet all of the regulation requirements [64]. For a product to be labelled "gluten-free," "no gluten," "free of gluten," or "without gluten," it must contain less than 20 parts per million of gluten. According to the FDA, this is the lowest level that can be reliably detected in foods using scientifically validated analytical methods [64]. Thus, even patients who believe they are strictly adhering to a GFD may ingest gluten on a daily basis.

According to the latest research, ingesting more than 50 mg of gluten per day causes intestinal damage for individuals with CeD. In order for damage to occur, the individual must eat at least five pounds of gluten-free food per day [11, 18]. Nonetheless, some patients with CeD need as little as 10 mg of gluten per day to promote small intestinal abnormalities.

A GFD is both socially troublesome and expensive. Gluten-free substitutes for items such as bread and pasta are more expensive than their gluten-containing counterparts. Additionally, research has shown compliance to the diet often to be poor as gluten contamination is difficult to avoid. For these reasons, many patients with CeD have expressed a desire for other alternative or complementary therapies that are less burdensome than a strict, lifelong GFD.

2.3.3 Diagnosis

The diagnosis of CeD has traditionally depended on intestinal biopsies alone. In current clinical practice, the diagnosis of CeD is based on a combination of clinical, serologic, and histologic factors. The initial test for diagnosis or exclusion of CeD is typically to assay for specific, serum, celiac-associated IgA or IgG antibodies, tissue transglutaminase-2, and deamidated gluten peptides. These serologic tests used alone, or in combination, are highly sensitive and specific for the diagnosis of CeD. However, these tests only work for patients who have not yet adopted the GFD. When a patient is maintaining a strict GFD without prior testing for CeD, evaluation for HLA-DQ2 and -DQ8 alleles can be useful as a negative result will largely exclude CeD [1].

In individuals suspected of having CeD, the American Gastroenterological Association mandates a biopsy to confirm the diagnosis. Although many still regard the small bowel biopsy as the "gold standard" in the diagnosis of CeD, the United States National Institutes

of Health released a consensus statement recommending biopsies only after a positive serology or when faced with incongruous or indeterminate serological results [22]. However, finding gluten-induced small intestinal mucosal injury is the hallmark of a CeD diagnosis. The mucosa will heal upon adaption of a GFD and the mucosal damage will reappear if gluten is reintroduced [36].

Small intestinal histologic abnormalities in CeD include villous atrophy, crypt hyperplasia, and lymphocytic infiltration of intraepithelial spaces. These findings are central to diagnosis and their severity over time are valuable to monitor disease course and response to therapy. Subjective methods to grade CeD histology severity include the Marsh-Oberhuber and Corazza-Villanacci systems. Quantitative histology uses villus height (Vh), crypt depth (Cd), and intraepithelial lymphocyte (IEL) count per 100 enterocytes to provide objective measures of histologic changes including the villus height-to-crypt depth (Vh:Cd) ratio [1]. These methods are discussed further in Section 2.3.6.

Most of the diagnostic features are expected to improve in the majority patients and generally normalize following the successful avoidance of gluten ingestion. Thus, favorable response to the GFD is an additional and important component of diagnosis; whereas failure to respond to starting a GFD is an indication for additional evaluation. Many physicians recommend repeat biopsy after six to 24 months of treatment with a GFD both to confirm the diagnosis of a gluten-sensitive enteropathy and to evaluate the degree of response to therapy [1].

While diagnosis can be difficult, it is imperative as undiagnosed, CeD can cause additional health complications including, but not limited to, type I diabetes, autoimmune thyroid disease, autoimmune liver disease, rheumatoid arthritis, and intestinal cancer. CeD can also likely cause malnutrition and vitamin deficiencies, leading to osteoporosis.

2.3.4 Small Intestinal Dynamics

The small intestinal epithelium is a single layer of cells organized into crypts and villi, as shown in Figure 2.1, known as the most rapidly self-renewing tissue in adult mammals. The cells that line the intestinal lumen perform the primary functions of digestion, water and nutrient absorption, and protection as a barrier against luminal pathogens [12].

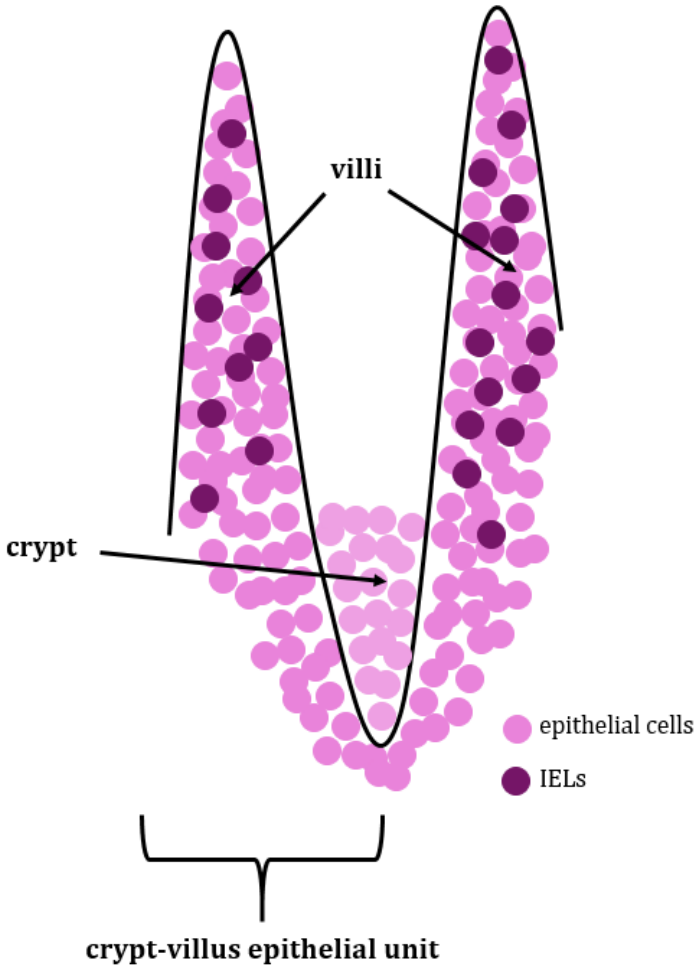

Figure 2.1: Illustration of the lumen of the small intestine. Villi are the finger-like protrusions into the gut lumen that help to absorb nutrients; crypts are the invaginations between villi. Villi and crypts are comprised of epithelial cells (represented by light purple circles). Intraepithelial lymphocytes (IELs, represented by dark purple circles) are T cells that exist in the lining of the intestine. The crypt-villus epithelial unit is comprised of a single villus and associated crypt. Figure is not drawn to scale.

The digestion of wheat, rye, and barley may expose the small intestine to immunoreactive epitopes that subsequently instigate a maladaptive immune response. Some foods remain undigested when entering the small intestine, and in turn contribute to the activation of specific populations of the immune system via T cells in the mucosa. Reactive T cell populations, in turn, mount a response targeting multiple autoantigens. The mechanism by which immunoreactive derivatives breach the mucosal epithelium is poorly understood. Possibilities include causing a reactive increase in the permeability of the tight junctions that line the small intestine, altering uptake and processing by enterocytes, or alternatively, the response may be instigated by activated IELs. The precise role of IELs in CeD is unclear, but IELs are thought to actively contribute the mucosal damage [22].

CeD primarily affects the superficial mucosa of the small intestine. The mucosal lesions vary in severity and extent, but the most characteristic features are a loss of normal villus structure with a reduction in villus height (Vh), an enlargement of the crypts with an increase in crypt depth (Cd), and inflammatory cell infiltration with an increase in the density of IELs [1].

In the small intestine, dividing cells are confined to the crypts, the invaginations of the epithelium into the underlying connective tissue between villi [12]. The intestinal stem cell is the crypt-based columnar cell that resides between Paneth cells at the crypt base. These stem cells divide both to maintain their own population at the base of the crypt and to produce proliferative transit-amplifying cells that migrate up the crypt which further divide and differentiate into terminally differentiated cell populations of the intestinal epithelium including the absorptive enterocytes, secretory goblet cells, enteroendocrine cells, and Paneth cells [7]. Stem cells are dangerous cells; they provide for renewal and repair, but disaster can ensue if they proliferate excessively or in the wrong place [20].

Adult stem cells at the base of the intestinal crypts proliferate and differentiate before migrating upward onto the villi, the relatively large finger-like protrusions into the gut lumen, where no further division occurs and all cells are fully differentiated. On the villi, the cells are exposed to the gut contents and shed from the villus tips into the gut lumen. The renewal process operates continually with cells taking two to seven days to make the journey from the site of their final division cycle in the crypt to the point of their exfoliation from the

villus tip [20]. In part because of current limitations impeding *in vivo* imaging of the entire crypt-villus epithelial unit (CVEU) over extended periods, the dynamics of cellular turnover in the gut are unclear for those suffering from CeD.

Villous atrophy is the blunting and flattening of the villi caused by the damage done by the immune system in a person with CeD after ingesting gluten. Damage to the villi can begin as early as three hours after exposure to gluten [11]. Figure 2.2 shows the structural changes to the small intestine due to villous atrophy and crypt hyperplasia, the over-filling of the crypts. Since intestinal villi are responsible for absorbing nutrients contained in food, losing them to villous atrophy can result in serious nutritional deficiencies. However, the villi are not permanently damaged as the intestines continuously renew themselves. For patients suffering from CeD, avoiding gluten should encourage mucosal healing.

Villi are not attacked by gluten. Gluten only triggers the reaction of the lymphocytes, which become "licensed to kill." The lymphocytes migrate into the lining of the gut and destroy the epithelial cells by producing chemicals that make them commit suicide, a process known as apoptosis. It takes a very short time for blunting to occur, often just a few days. The amount of time that it takes the villi to heal themselves is highly variable. It can take anywhere from one to two weeks to several years. Unfortunately, complete healing never occurs for some [11]. A patient is not guaranteed to have CeD just because they have villous atrophy. Several other conditions, plus some medications and bacterial overgrowth, can destroy intestinal villi.

2.3.5 Intraepithelial Lymphocytes

Lymphocytes comprise about 25% of a person's white blood cells and include the B cells that mature in bone marrow and T cells that mature in the thymus, each playing key roles in the development of immunity. Intraepithelial lymphocytes (IELs) are T cells that exist in the lining (intraepithelil) of the intestine. IELs are located between epithelial cells of the small intestinal villus and crypt [32].

Cellular and molecular cross-talk between epithelial cells and IELs in the gut epithelium appears to play a key role in the reciprocal growth and activation of these cells and in the maintenance of intestinal homeostasis. Interferon gamma and other cytokines (e.g.,

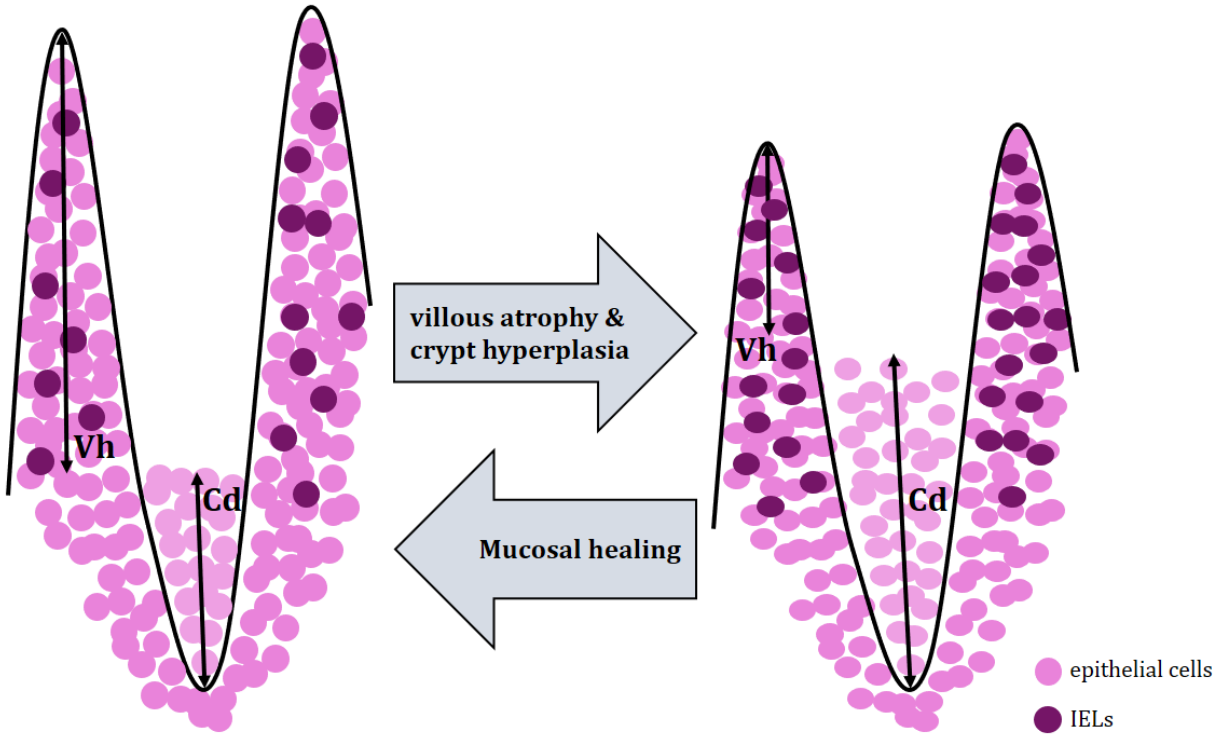

Figure 2.2: Illustration of small intestinal mucosal sections for patients with CeD before and after the ingestion of gluten. Adapted from [1] to illustrate the mucosal sections cut perpendicular to the luminal surface for (left) a patient with well-treated CeD on a GFD with labelled villus height (Vh) and crypt depth (Cd) and (right) the changes to these measurements for a patient with incompletely-treated CeD. The light purple circles represent the epithelial cells that comprise the small intestinal mucosa. The dark purple circles depict the increased levels of intraepithelial lymphocytes (IELs). Figure is not drawn to scale.

Interleukin-2, Interleukin-4, tumor necrosis factor α, and transforming growth factor β) potentially secreted by IELs modulate barrier function of the epithelium and intestinal epithelial cell functions such as cytokine production [32].

IELs are anatomically exposed to luminal antigens and pathogens, which continuously interact with and potentially activate these cells. This notion is supported by the potential of IELs to constitutively secrete cytokines and to display cytotoxic activity. Increased activation of IELs occurs in the small intestinal villus tip compared with the crypt regions. In fact, epithelial villus tips are enriched in mature epithelial cells, which display numerous biological functions and thus may facilitate interactions between IELs and luminal antigens, which account for the higher activation state [32].

In patients suffering from untreated CeD, IEL densities per 100 enterocytes is often increased. Although this density is increased, the total number of IELs may not be increased since the absorptive surface area is markedly reduced.

2.3.6 Q-MARSH Score

Quantitative-Mucosal Algorithmic Rules for Scoring Histology (Q-MARSH), as described by Adelman et al. [1], aligns quantitative histology results with the traditional Marsh, Marsh-Oberhuber, and Corazza-Villanacci systems as used in the diagnosis of CeD. The Marsh classification was originally developed to help describe associations between gluten exposure, polymorphisms of the major histocompatibility complex, and mucosal pathology representative of the spectrum of mucosal appearances observed in response to gluten exposure. The classification was not intended for use in clinical practice. Marsh originally described five interrelated lesions: pre-infilrative, infiltrative, hyperplastic, destructive, and hypoplastic [1].

This classification was later modified by Oberhuber, named the Marsh-Oberhuber classification, and was intended to be used by clinical pathologists to assist in diagnosing CeD through assessing and categorizing the degrees of mucosal abnormality in small intestinal biopsies. The classification of one of six diagnostic grades as described in Table 2.1 relies on the subjective judgement of the pathologist and shows low interobserver agreement and reduced reproduciblity [1].

Table 2.1: Marsh-Oberhuber histologic grading in CeD [1].

Grade	Description	Components
0	Normal	normal appearing villus architecture
I	Infiltrative	normal mucosal and villus architecture; increased number of IELs
II	Hyperplastic	similar to I with enlarged crypts and increased crypt cell division
IIIa	Partial villous atrophy	shortened, blunt villi; mild lymphocyte infiltration, enlarged hyperplastic crypts
IIIb	Subtotal villous atrophy	clearly atrophic villi, enlarged crypts whose immature epithelial cells are generated at an increased rate, influx of inflammatory cells
IIIc	Hypoplastic	total villous atrophy, complete loss of villi, severe crypt hyperplasia, infiltrative inflammatory lesion

Another classification, the Corazza-Villanacci score, divides the spectrum of CeD into three grades. Grade A (architecturally intact), so called non-atrophic, includes normal crypts and villus architecture, but increased IELs (utilizing a threshold of 25 IELs per 100 enterocytes). Grade B1 (partial villous atrophy) is atrophic with Vh:Cd ratio of less than 3.0, but where villi are still clearly detectable and IELs are increased. Grade B2 (total villous atrophy) is atrophic to the extent that villi are flat and not detectable, also including crypt hyperplasia and increased IELs [1].

The three key elements that are commonly noted as a part of a routine, clinical pathology reporting include villus height (Vh), which is often referred to in terms of a reduction (e.g., villus "blunting," "flattening," or "atrophy"); crypt depth (Cd), which is often referred to in terms of enlargement (e.g., "elongation" or "hyperplasia"); and an increase in intraepithelial lymphocyte (IEL) density, especially at the villus tip. The primary architectural changes seen in CeD, reduced Vh and increased Cd, as well as the most characteristic element of inflammatory cell infiltration (increased IEL density), are used to obtain objective, continuous measurements of the severity of the celiac lesion in a particular biopsy [1].

Standardized procedures to perform quantitative histology include measurements of Vh and Cd from five to 12 crypt-villus epithelial units (CVEUs) from at least four biopsy specimens taken from the distal portion of the second or third part of the duodenum. A light microscope with a calibrated micrometer is used to measure individual Vh and Cd, in μm, from a well-oriented CVEU. The ratio of Vh to Cd is commonly used as a single value, the villus height-to-crypt depth ratio (Vh:Cd), to encompass both elements of mucosal architectural change. The Vh:Cd of each CVEU is calculated and the reported ratio is derived from the average of these values. This differs from the traditional approach of scoring the single worst area of the biopsy when using histology to confirm a diagnosis of CeD. Because measurements of Vh and Cd are obtained independently and each can be used as an individual measure, as well as the quantification of IELs per 100 enterocytes, quantitative histology provides three distinct, quantifiable, and relevant measures of small bowl mucosal change over time in response to therapeutic intervention [1].

Q-MARSH is a quantitative refinement of the qualitative and categorical Marsh, Marsh-Oberhuber, and Corazza-Villanucci Scales, as described and compared in Table 2.2. Both

Table 2.2: Q-MARSH compared with Marsh, Marsh-Oberhuber, and Corazza-Villanacci, as introduced in Adelman et al. [1].
Key:
* Varies, but approximately 30 per 100 enterocytes.
** Crypt hyperplasia or elongation.
*** Villus atrophy.
† Increased villous atrophy.
‡ Partial villous atrophy.
†† Complete villous atrophy.

Q-MARSH		**Marsh**	**Marsh-Oberhuber**	**Corazza-Villanacci**
IELs	**Vh:Cd**			
Normal	Vh:Cd $\geq$ 3.0	0	0	Normal
Increased*	Vh:Cd $\geq$ 3.0	1	1	Grade A
Increased	2.0 $\leq$ Vh:Cd $<$ 3.0	2**	2**	
Increased	1.0 $\leq$ Vh:Cd $<$ 2.0	3***	3a †	Grade B1 ‡
Increased	0.5 $\leq$ Vh:Cd $<$ 1.0		3b †	
Increased	Vh:Cd $<$ 0.5		3c †	Grade B2 ††

Q-MARSH and Marsh-Oberhuber take into consideration the key elements of mucosal architecture (i.e., Vh, Cd, and IEL density); however, the Marsh-Oberhuber classification incorporates subjective, non-quantitative evaluation of the key mucosal elements into an assigned categorical grading of the appearance of the mucosa. In contrast, quantitative histology and methodology relies on the objective, quantitative, and continuous measures of the key mucosal structural elements to assess changes over time [1].

2.4 Related Mathematical Models

To our knowledge, only three mathematical models of celiac disease have been developed. Trask [63] presented a model that extended a similar model by Forde and Meeker [24]. Both mathematical models used a system of ODEs to quantify the interactions in a patient with CeD causing the small intestinal epithelium to become permeable following the ingestion of gluten. The third mathematical model also utilized a system of ODEs was developed by Demin et al. [21] and focused on predicting the efficacy of drug treatments for CeD and included detail necessary to test the effectiveness of these particular treatments. Details of each of these models will be discussed further. We also introduce a mathematical model by Parker et al. [50] which focused on cell tracking between villi and crypts using a system of ODEs.

Trask [63] aimed to use a mathematical model of ODEs to better understand the biological mechanisms in CeD. Her work addressed the way in which gluten induces zonulin, a protein known to regulate the permeability of the intestine, production in the small intestine for a patient suffering from CeD. Since gliadins are the constituent of gluten to which patients with CeD react, gliadins were used to represent gluten in her model. Her five-dimensional system included gliadin on both the apical and basolateral sides (inside and outside, respectively) of the small intestine, zonulin, the level of permeability, and the immune response, represented by T cell concentration. Simulations included those for a healthy subject and patient with CeD, as well as for a patient with CeD ingesting gluten multiple times per day.

Stability analysis of her model indicated that there was the same steady state regardless of the level of permeability and that the permeability will relax to normal in the absence of gluten. Thus, in her model, a GFD is always effective.

Trask's model was focused on understanding the permeability of the gut and tracking gluten as it passes through the lining of the small intestine. In order to build upon this model, more data would be needed to quantify the rates at which gluten can pass through the intestinal wall as well as the permeability of the gut mucosa dependent on zonulin levels. At the present time, clinical studies pertaining to CeD do not collect this data, making accurate parameterization difficult.

Demin et al. [21] developed a mathematical model of the immune response in CeD with a focus on predicting the efficacy of a tissue transglutaminase-2 inhibitor as well as other potential drug treatments for CeD by analyzing the intestinal villous area and antibody levels before and after treatment. Their model accounted for the complex regulatory mechanisms associated with CeD by including both the innate and adaptive immune responses.

To describe these processes, Demin et al. [21] used a system of ODEs including 16 variables and 54 parameters. Of the parameters, 17 were estimated from literature data, two were assumed, four were calculated using experimental data, and 31 were evaluated based on the best fitting to the appropriate experimental data. The 16 variables included gluten in both the lumen and lamina, intestinal epithelial cells, IELs, zonulin, cytokines, antigen presenting cells, and antibodies. To describe healthy patients, they assumed that there are no innate and adaptive immune responses in healthy subjects. This approach allowed a simplification of the equations in the model and decreased the number of model parameters.

Their mathematical model was focused on describing the mechanisms of IEL activation and better understanding the processes underlying the absence of immune responses in healthy subjects. The level of detail was included to very accurately describe the immune response in order to test particular therapies. Their model could not be linked to levels of gut health and the Vh:Cd.

While not a mathematical model applied to CeD, Parker et al. [50] established a cell-tracking method to collect data about the distribution of labeled cells along the CVEU over time and developed a mathematical model to quantify crypt cell production and villus cell

migration in the proximal and distal small intestine of mice under both normal conditions and when proliferation is reduced and temporarily halted. The purpose of their study was to investigate whether cell proliferation within the small intestinal crypts is sufficient to explain the observed cell migration on villi, during homeostasis and under altered conditions.

The time-dependent mathematical model of the CVEU distinguishes between two physical compartments: the crypt compartment, which includes all proliferative cells in the CVEU as well as some nonproliferative cells, and the villus compartment, which contains all remaining nonproliferative cells. The CVEU was defined as a one-dimensional column of cells running from the base of a crypt to the tip of a neighboring villus. Their model assumes that cell transfer between the crypt and villus is exclusively related to cell proliferation. This assumption derives from experimental conditions in which the size of the crypt compartment maintains at a constant fixed value over time. Cell proliferation within the crypt is compensated for by the transfer of cells to the villus, and these processes take place at the same rate. Given that the CVEU is one-dimensional, the size of the crypt compartment is also the length of the crypt, expressed by the number of cells. Shedding of cells into the gut lumen does not begin until cells reach the tip of the villus [50].

Migration of labeled cells from the crypt to the villus is initiated when the number of labeled cells within the crypt reaches threshold value. Similarly, cell shedding from the villus is initiated when the number of labeled cells in the villus reaches their threshold value. Their model uses the Heaviside function so that cell transfer occurs only if the compartment has reached its threshold. Parker et al. [50] collected data from the small intestines of mice. The observed counts of labeled cells in both the crypt and villus was used to estimate parameters in the mathematical model.

The focus of their work was to determine the relationship between cell proliferation and cell migration which drove many of the assumptions, including the size of the crypt compartment. In CeD, the lengths of the crypts are elongated in a process called crypt hyperplasia. In a mathematical model to understand the dynamics of CeD, the size of the crypt compartment cannot remain fixed if the model is to accurately represent crypt hyperplasia. We use the structure from Parker et al. [50] to describe cellular transfer.

2.5 Celiac Disease Model

We aim to develop a mathematical model to assist in uncovering the dynamics of cellular turnover involved in inflammation and healing in the small intestine as related to celiac disease and other comparable diseases termed "leaky gut." Our mechanistic mathematical model replicates the Q-MARSH diagnostic method described in Section 2.3.6 for both healthy individuals and patients suffering from CeD to better understand the immunological dynamics of villus atrophy and crypt hyperplasia in CeD. The model was integrated with experimental data and parameters, both from literature, to ensure outputs were compatible with data. By looking at the changes in the model under various activation functions of the immune system, we hope to investigate the progression of this disease as well as improve diagnostic methods and suggest potential therapies to mitigate the damaging effects of CeD.

Studies have provided evidence that following gluten ingestion, an inflammatory process with dose-dependent accumulation of IELs is followed by mucosal villous atrophy with crypt hyperplasia [36]. However, the only method to determine the extent of this intestinal damage is via an upper endoscopy and biopsy, a rather invasive procedure. Through simulation of the time courses of gut health, it may be possible to better understand and determine methods to mitigate this damage. While patients with untreated CeD exhibit a variety of symptoms, the presence of symptoms is not related to the histological features of a biopsy [36].

Before the development of a mathematical model, assumptions regarding the dynamics of the cellular populations must be made. First and foremost, there should be no immune response for healthy subjects or patients with well-treated CeD on a GFD. When a patient with CeD responds positively to and maintains a GFD, the small intestine will heal and return to a "normal" state. Thus, in the absence of immune activation, there should be no change in the cellular populations.

Further, to simplify the model, we assume that the IELs are the source of damage to the villi and elongation of the crypts. While there are intermediate factors causing the small intestinal damage, the relation between the IELs, these factors, and the cellular populations is neither directly known nor easily quantifiable. We assume that the increase in the populations causing the intestinal damage is proportional to the IEL population. Thus,

in our model, IELs represent the immune response and are the source of damage to the small intestinal lining.

2.5.1 Model Formulation

So that our model may be applied to other small intestinal disorders, gluten is not explicitly represented in our system. Instead, we use a function, $A(t)$, to represent the activation of the immune response. In a patient suffering from CeD, this activation can be related to gluten intake and the amount of gluten in the small intestine. Multiple forms of immune activation are considered in Section 2.6.2.

We mathematically model the dynamics of cellular turnover involved in inflammation and healing in the small intestine with a focus on the two diagnostic factors involved in Q-MARSH described in Section 2.3.6, the villus height-to-crypt depth ratio (Vh:Cd) and number of IELs. In order to describe the dynamics of the Vh:Cd, we consider the dynamics of the number of villus cells at both the base and tip of the villi, V_1 and V_2, respectively, and the number of cells in the crypts, C. We assume that the total number of cells in the villus and crypt are proportional to the villus height, Vh, and crypt depth, Cd, respectively. Since all small intestinal mucosa have a density of IELs, we consider the additional IELs that are activated following an immune response. The dynamics of the activated IELs, I, are considered directly.

Our model can be described using the system of ODEs given by

$$\begin{cases} \frac{dV_1}{dt} &= r_1 C \cdot H\left(1 - \frac{V_1}{K_1}\right) - r_2 V_1 \cdot H\left(1 - \frac{V_2}{K_2}\right) - d_1 V_1 I \\ \frac{dV_2}{dt} &= r_2 V_1 \cdot H\left(1 - \frac{V_2}{K_2}\right) - d_V V_2 - d_2 V_2 I \\ \frac{dC}{dt} &= r_C C(1 - \frac{C}{K_C}) + pCI - r_1 C \cdot H\left(1 - \frac{V_1}{K_1}\right) \\ \frac{dI}{dt} &= A(t) - \gamma I \\ \text{Vh} &= \alpha(V_1 + V_2) \\ \text{Cd} &= \beta C \end{cases} \tag{2.1}$$

where

$$H(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \tag{2.2}$$

is the Heaviside function. The use of the Heaviside function allows a sharp turn on and off of cellular transfer when cells reach their transfer threshold as used in [50]. The compartments of the model are described in Table 2.3 and the dynamics of the model are described by Figure 2.3.

In our model, the activation function $A(t)$ is the only source for IELs. In the absence of activation, the number of IELs decays exponentially at rate γ. In the case of a healthy individual, $A(t) \equiv 0$. We consider different forms of $A(t)$ to understand how activation of the immune response changes both short- and long-term behavior in Section 2.6.2.

Although many crypts contribute to a single villus, we consider single crypt-villus epithelial units similar to other related mathematical models [50]. A CVEU is defined to be a single continuous strip of epithelial cells running from the base of a particular crypt to the tip of the associated villus.

Cells in the villus are broken into two compartments, those at the base, V_1, and tip, V_2. The sum of both compartments gives the total number of cells in the villus, V. The cells at the base of the villus grow exponentially at rate r_1, proportional to the number of cells in the crypts. Stem cells regenerate in the crypts which migrate to form the columnar cells in the villus. As such, this is the only source of villus cells. We assume that the cells move from the crypt to the base of the villus linearly; however, this transfer is turned off with the use of Heaviside function (2.2) when V_1 cells reach their transfer threshold, K_1. Similarly, V_2 cells mature from V_1 cells at rate r_2, with a transfer threshold of V_2 cells, K_2. Whenever V_1 and V_2 exceed their respective transfer thresholds, these cellular transfers are turned off and cells remain in their current state.

Villus cells were split into two compartments as only the cells at the tip of the villus shed at rate d_V into the gut lumen during cell turnover. However, the activated IELs reduce both types of villus cells as the immune system mistakenly attacks the small intestine. We use two different rates of destruction, d_1 and d_2, as IELs are more dense in the tip of the villus and therefore have a greater effect. Thus, $d_1 \leq d_2$.

Table 2.3: Compartments of CeD model.

Compartment	Description	Units
V	total cells in the villus ($V_1 + V_2$)	cells
V_1	cells in base of the villus	cells
V_2	cells in tip of the villus	cells
C	cells in the crypt	cells
I	activated intraepithelial lymphocytes (IELs)	IELs
Vh	villus height	μm
Cd	crypt depth	μm
Vh:Cd	villus height-to-crypt depth ratio	dimensionless

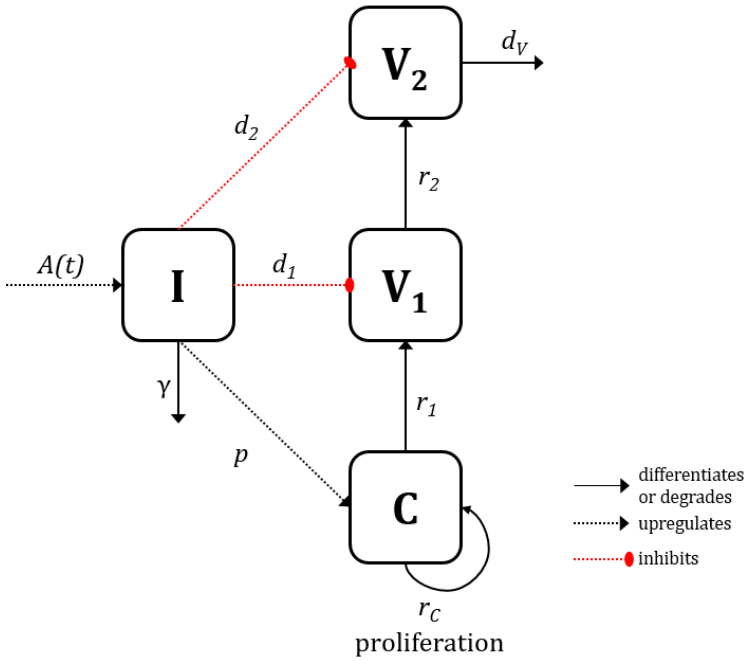

Figure 2.3: Schematic of model formulated to understand the dynamics of cellular turnover during the immune response of CeD.

Even in the presence of IELs, the small intestine will attempt to heal itself. Cells in the crypt grow logistically at rate r_C to carrying capacity, K_C. However, the presence of IELs causes the crypt cells to proliferate at an increased rate, p. This increase in production causes a "spill-over" effect. Thus, when the immune system is activated, we will see the crypts elongate during crypt hyperplasia.

In order to replicate the Q-MARSH score, we relate villus and crypt cells to the villus height and crypt depth, respectively. We assume that the respective cells are proportional to Vh and Cd with length conversion factors α and β, respectively. All parameter values are discussed in Section 2.6.1.

It is known that when a patient with CeD ingests gluten, a chain of events leading to an immune response is initiated and mediated by lymphocytes. The villi are not attacked by the gluten; rather, gluten triggers the reaction of the lymphocytes which then become "licensed to kill," migrate into the lining of the gut, and destroy the cells of the lining by producing chemicals that make them commit suicide, a process known as apoptosis [11]. Only IELs cause damage; the presence of gluten is the initiator of the immune response. As the cellular dynamics in both the villus and crypt change, the Vh:Cd also changes. In the Q-MARSH, both the Vh:Cd and level of IELs are considered to diagnosis a patient with CeD. After confirmation that this simplified model accurately depicts the histological changes, more detail can be made by including other key players such as cytokines, autoantigens, and enzymes. Through understanding these dynamics, the goal is to improve diagnostic methods and find therapies that limit the destruction of the small intestine.

2.5.2 Existence and Uniqueness of Solutions

We begin by proving existence, uniqueness, and non-negativity of solutions to our system (2.1) for a given initial condition vector in $\mathbb{R}^4_+$. Assuming $A(t)$ is in $L^1[0,T]$, the solution of the $I(t)$ differential equation exists for all $t > 0$ by explicitly integrating the differential equation. Additionally, this component is uniformly bounded for all $t \geq 0$. We find that

$$\frac{dI}{dt} = A(t) - \gamma I \tag{2.3}$$

$$\implies I(t)e^{\gamma t} - I(0) = \int_0^t e^{\gamma\tau} A(\tau) d\tau \tag{2.4}$$

$$\implies I(t) = \left(I_0 + \int_0^t e^{\gamma\tau} A(\tau) d\tau \right) e^{-\gamma t}. \tag{2.5}$$

Assuming $A(t)$ is given, consider the system with V_1, V_2, and C. We write our reduced system as

$$\begin{aligned} x'(t) &= f(t, x(t)) \\ x(t_0) &= \phi, \end{aligned} \tag{2.6}$$

where $f : N \to \mathbb{R}^3$, $N \subset [0, T] \times \mathbb{R}^3$ is open, and $\phi \in \mathbb{R}^3$. We define R to be the rectangular region

$$R = R((t_0, \phi), a, b) = \{(t, x) : |t - t_0| \leq a, |x - \phi| \leq b\}, \tag{2.7}$$

where a and b are positive numbers.

For b sufficiently small, $V_{1_0} < K_1$, and $V_{2_0} < K_2$, the function f is continuous on x for fixed t and measurable in t for a fixed x, and uniformly bounded on the region R. With these initial conditions, the Heaviside functions are constant at 1. We can obtain solutions locally in time near these initial conditions, using results from Lukes [42]. The structure of the system of differential equations will give uniform L^∞ bounds and one can use Lipschitz properties of the right hand side of the differential equation system to get uniqueness of the solutions.

The next theorem, from [57], specifies conditions on f which ensure that non-negative initial conditions will guarantee non-negative solutions to initial value problem (2.6). We denote the right-hand side of the system (2.1) as $f_1, \ldots, f_4$.

Theorem 2.1. *Assume that when $\phi \in \mathbb{R}^n$ satisfies $\phi \geq 0$ with $\phi_i(0) = 0$ for some i, then $f_i(\phi) \geq 0$. Then, if $\phi \in \mathbb{R}^n$ satisfies $\phi \geq 0$, the solution of (2.6) satisfies $x(t) \geq 0$ for all $t \geq t_0$.*

We first use Theorem 2.1 to prove that non-negative initial conditions will guarantee non-negative solutions to our system (2.1), assuming a solution exists. This theorem requires f

is continuous, so we need to be in the region where the Heaviside function is equal to 1. We will then show existence and uniqueness.

Theorem 2.2. *For any vector of initial conditions* $\phi = (V_{1_0}, V_{2_0}, C_0, I_0) \geq 0$ *with* $V_{1_0} < K_1$ *and* $V_{2_0} < K_2$*, and* $A(t) \geq 0$*, the system (2.1) has a non-negative solution.*

Proof. We begin by checking the hypotheses of Theorem 2.1.

1. Assume $\phi = (0, V_{2_0}, C_0, I_0) \geq 0$. Then

$$f_1(\phi) = r_1 C_0 \geq 0.$$

2. Assume $\phi = (V_{1_0}, 0, C_0, I_0) \geq 0$. Then

$$f_2(\phi) = r_2 V_{1_0} \geq 0.$$

3. Assume $\phi = (V_{1_0}, V_{2_0}, 0, I_0) \geq 0$. Then

$$f_3(\phi) = 0 \geq 0.$$

4. Assume $\phi = (V_{1_0}, V_{2_0}, C_0, 0) \geq 0$. Then

$$f_4(\phi) = A(t) \geq 0.$$

Therefore, whenever $\phi \geq 0$ (with $\phi_i(0) = 0$ for some i), we have $f_i(\phi) \geq 0$. Thus, by Theorem 2.1, if a solution to our system (2.1) exists for $\phi \geq 0$, the solution is non-negative for all $t \geq 0$. □

Now, we wish to show that our system (2.1) is also bounded from above. Thus, if a solution exists, then our states are bounded. We require that

$$\int_0^\infty A(t)dt < \infty \tag{2.8}$$

for all choices of $A(t)$ and that $A(t)$ is bounded above by $M_1 \in \mathbb{R}$. Thus, we can further bound $I(t)$ given in equation (2.5).

$$I(t) \leq \left(I_0 + \int_0^t e^{\gamma\tau} M_1 d\tau\right) e^{-\gamma t} \tag{2.9}$$

$$= \left(I_0 + M_1 \frac{e^{\gamma t} - 1}{\gamma}\right) e^{-\gamma t} \tag{2.10}$$

$$= I_0 e^{-\gamma t} + \frac{M_1}{\gamma}(1 - e^{-\gamma t}) \tag{2.11}$$

which is bounded above by some constant $M_2 \in \mathbb{R}$. Therefore, we have that $I(t) \leq M_2$ for all time.

Next, we consider supersolutions for the remaining differential equations. Each supersolution will have the same initial condition as the respective differential equation, but will remain above the differential equation for all time. We will denote the supersolutions with the lowercase letters of their corresponding differential equations. For example, our supersolution, $c(t)$, is greater than $C(t)$ for all time.

We have that

$$\frac{dC}{dt} = r_C C(1 - \frac{C}{K_C}) + pCI - r_1 C \cdot H\left(1 - \frac{V_1}{K_1}\right).$$

Since $r_1, C, H(1 - \frac{V_1}{K_1}) \geq 0$, we know that the last term of $\frac{dC}{dt}$ will pull down $C(t)$ and can be dropped from the differential equation for the supersolution. Further, since $I(t) \leq M_2$, the derivative of our super solution can be written

$$\frac{dc}{dt} = r_C c(1 - \frac{c}{K_C}) + pcM_2 = c\left(r_C + pM_2 - \frac{r_C c}{K_C}\right). \tag{2.12}$$

Recall that $c(t)$ decreases when $\frac{dc}{dt} < 0$ which occurs when

$$r_C + pM_2 - \frac{r_C c}{K_C} < 0$$

since $0 \leq C \leq c$. Rearranging, this is the same as

$$c > K_C \left(1 + \frac{pM_2}{r_C}\right). \tag{2.13}$$

Since $c(t)$ decreases when inequality (2.13) holds, the upper bound for our supersolution $c(t)$ is $\max\left(C_0, K_C(1 + \frac{pM_2}{r_C})\right)$. Since $c(t)$ is the supersolution for $C(t)$ for all t, we have that $C(t)$ is bounded above by $\max\left(C_0, K_C(1 + \frac{pM_2}{r_C})\right)$.

Now, we consider the supersolution for

$$\frac{dV_1}{dt} = r_1 C \cdot H\left(1 - \frac{V_1}{K_1}\right) - r_2 V_1 \cdot H\left(1 - \frac{V_2}{K_2}\right) - d_1 V_1 I.$$

We see that since $r_2, d_1, V_1, I, H(1 - \frac{V_2}{K_2}) \geq 0$, the last two terms will pull down $V_1(t)$. Thus, we consider the supersolution to

$$\frac{dv_1}{dt} = r_1 C \cdot H\left(1 - \frac{v_1}{K_1}\right). \tag{2.14}$$

When $v_1 \geq K_1$, we have that

$$H(1 - \frac{v_1}{K_1}) = 0 \tag{2.15}$$

$$\implies \frac{dv_1}{dt} = 0 \tag{2.16}$$

$$\implies v_1(t) \equiv k \text{ for } k \in \mathbb{R}. \tag{2.17}$$

When $v_1 < K_1$ and we use supersolution $c(t)$, we have $\frac{dv_1}{dt} = r_1 c$. Therefore, we will have that $v_1(t)$ will increase to $v_1 = K_1$ since once we hit K_1, $v_1(t) \equiv k$. Thus, the upper bound on v_1 is $\max(V_{1_0}, K_1)$ since if $V_{1_0} \geq K_1$, $v_1(t)$ will remain at V_{1_0} for all t. Thus, $V_1(t)$ is bounded above by $\max(V_{1_0}, K_1)$.

Similarly, we see for

$$\frac{dV_2}{dt} = r_2 V_1 \cdot H\left(1 - \frac{V_2}{K_2}\right) - d_V V_2 - d_2 V_2 I,$$

we can solve

$$\frac{dv_2}{dt} = r_2 v_1 \cdot H\left(1 - \frac{v_2}{K_2}\right) \tag{2.18}$$

for our supersolution. For the similar reasons as $v_1(t)$, we have that the upper bound on v_2 is $\max(V_{2_0}, K_2)$ and $V_2(t)$ is bounded above by this maximum.

Since $v_1(t), v_2(t)$, and $c(t)$ are all supersolutions, they will always remain above our solutions $V_1(t), V_2(t)$, and $C(t)$, provided these solutions exist. Combining with our results from Theorem 2.2, we have

$$0 \leq V_1(t) \leq \max(V_1(0), K_1), \tag{2.19}$$

$$0 \leq V_2(t) \leq \max(V_2(0), K_2), \tag{2.20}$$

$$0 \leq C(t) \leq \max\left(C(0), K_C(1 + \frac{pM_2}{r_C})\right), \text{ and} \tag{2.21}$$

$$0 \leq I(t) \leq M_2. \tag{2.22}$$

2.5.3 Long-Term Solutions and Stability Analysis

We now find the equilibrium points of our model and their stability in the absence of an immune response. We begin by noting that when there is no immune response, our system represents a healthy small intestinal structure with neither villous atrophy nor crypt hyperplasia. That is, when $I(t) \equiv 0$ and $A(t) \equiv 0$, we expect V_1, V_2, and C to remain at their initial conditions V_{1_0}, V_{2_0}, and C_0, respectively, since there will be no cellular turnover. Additionally, since we are considering a healthy subject, we assume that the initial conditions for these the villi classes are below their respective transition thresholds, that is, $V_{1_0} < K_1$ and $V_{2_0} < K_2$. In our model, this is representative of the Heaviside functions being "on" at the initial time. When $I(t) \equiv 0$, $A(t) \equiv 0$, and $H(1 - \frac{V_i}{K_i}) = 1$ for $i = 1, 2$, our model in equation (2.1) simplifies to

$$\begin{cases} \frac{dV_1}{dt} &= r_1 C - r_2 V_1 \\ \frac{dV_2}{dt} &= r_2 V_1 - d_V V_2 \\ \frac{dC}{dt} &= r_C C(1 - \frac{C}{K_C}) - r_1 C \end{cases} \tag{2.23}$$

Since we are searching for equilibrium points, we set V_1', V_2', and C' equal to 0 and solve for V_1, V_2, and C. Solving this system, we have two potential equilibrium points $(0,0,0)$, the trivial equilibrium point, and $(\frac{r_1K_C}{r_2}(1-\frac{r_1}{r_C}), \frac{r_1K_C}{d_V}(1-\frac{r_1}{r_C}), K_C(1-\frac{r_1}{r_C}))$. For the second equilibrium point to be biologically feasible, we need $1-\frac{r_1}{r_C} > 0$, or $r_1 < r_C$, and will choose this as an assumption for our model.

To determine the stability of these equilibrium points, we first find the Jacobian matrix in the absence of the immune response and when $V_1 < K_1$ and $V_2 < K_2$ for all t. This can be represented as

$$\begin{bmatrix} -r_2 & 0 & r_1 \\ r_2 & -d_V & 0 \\ 0 & 0 & r_C - \frac{2r_CC}{K_C} - r_1 \end{bmatrix}$$

Next, we evaluate the Jacobian at both of our equilibrium points and find the eigenvalues associated with each matrix. The sign of our eigenvalues determines the stability of the associated equilibrium point.

1. When we evaluate the Jacobian at $(0,0,0)$, the trivial equilibrium point, we have

$$\begin{bmatrix} -r_2 & 0 & r_1 \\ r_2 & -d_V & 0 \\ 0 & 0 & r_C - r_1 \end{bmatrix}$$

   which gives eigenvalues $-r_2$, $-d_V$, and $r_C - r_1$. Since $r_C - r_1 > 0$, this equilibrium point is unstable.

2. When we evaluate the Jacobian at $(\frac{r_1K_C}{r_2}(1-\frac{r_1}{r_C}), \frac{r_1K_C}{d_V}(1-\frac{r_1}{r_C}), K_C(1-\frac{r_1}{r_C}))$, we have

$$\begin{bmatrix} -r_2 & 0 & r_1 \\ r_2 & -d_V & 0 \\ 0 & 0 & r_1 - r_C \end{bmatrix}$$

   which has eigenvalues $-r_2$, $-d_V$, and $r_1 - r_C$. Since $r_1 < r_C$, all three eigenvalues are negative and this equilibrium point is stable.

Since the second equilibrium point is stable, we interpret this as a healthy subject in our system. Also note that since we want $V_1 < K_1$ and $V_2 < K_2$ on our time interval, we require $\frac{r_1 K_C}{r_2}(1 - \frac{r_1}{r_C}) < K_1$ and $\frac{r_1 K_C}{d_V}(1 - \frac{r_1}{r_C}) < K_2$.

2.6 Numerical Simulations

To simulate the model for both a healthy subject and a patient with celiac disease, we first determine feasible ranges for our parameters using the literature. Using these ranges, we can then ensure that our model can simulate a healthy subject without an immune response. In these simulations, we will see that there is no cellular turnover due to the lack of immune response. We will then include the immune response to simulate a patient with CeD for two scenarios - a single, transient immune activation and a pulsed, daily activation recreating a gluten challenge. In order to produce these simulations, parameters involved in the immune response will need to be hypothesized as they cannot be found in the literature. Finally, we want to ensure that our model can reproduce patient-level data and narrow down feasible ranges for those parameters that could not be identified in the literature using data from a gluten-challenge study.

2.6.1 Parameter Selection

Our working model has 16 parameters, 11 of which are included in the system of ODEs, two to relate cellular counts to diagnostic measurements, and three in the source function for IELs. Of these values, ten were either found in or estimated from literature; the remaining six were estimated numerically using data from literature and model output. A full list of parameters and their descriptions, units, and values are given in Table 2.4. Since villi and crypt cells are both epithelial cells, they have the same units of cells. We assume that both V_1 and V_2 cells are absorptive enterocytes, also known as columnar cells. For the V_1 cells, we will use transition rates for transit-amplifying cells, even though transit-amplifying cells are generally found at the top of the crypt; while for V_2 cells, we will use the transition rates for enterocytes.

Table 2.4: Parameter values and their units for CeD model.
Key: * Current assumption. ** Calculated in Section 2.6.2.

Parameter	Description [Units]	Healthy Subject	Celiac Patient	Source
r_1	cell differentiation rate from crypt to villus base [1/min]	$0.0010 - 0.0017$	$0.0010 - 0.017$	[7, 50]
r_2	cell migration rate from villus base to tip [1/min]	$0.0021 - 0.0035$	$0.00069 - 0.0014$	[22]
K_1	transition threshold for base villus cells [cells]	$921 - 969$	$748 - 787$	[7, 20]
K_2	transition threshold for tip villus cells [cells]	$2,490 - 2,620$	$2,024 - 2,129$	[7, 20]
d_V	villus cell turnover rate [1/min]	$0.00025 - 0.00026$	$0.00025 - 0.00026$	[7, 21]
d_1	destruction rate of villus cells in the base by IELs [1/(min·IELs)]	—	0	*
d_2	destruction rate of villus cells in the tip by IELs [1/(min·IELs)]	—	$0.00001 - 0.00052$	**
r_C	(healthy) cell proliferation rate in crypt [1/min]	$0.020 - 0.031$	$0.020 - 0.031$	[7, 50]
K_C	(healthy) carrying capacity of crypt cells [cells]	$244 - 256$	$244 - 325$	[20]
p	(additional) proliferation rate of crypt cells due to IELs [1/(min·IELs)]	—	$0.00010 - 0.0071$	**
γ	(natural) decay rate of IELs [1/min]	—	$0.00091 - 0.00092$	[21]

The Marsh Classification used in the diagnosis of CeD considers IEL counts in both the jejunumn and duodenum. However, since the biopsy to confirm CeD is generally taken from the duodenum, we consider the cellular turnover and activated IEL counts in the duodenum when searching for appropriate parameters. Due to the invasiveness of human biopsies, there is little data on the cellular rates in the small intestine in humans. However, there are more studies and data as related to CeD and the small intestine in the mouse. We assume rates for cellular transition in the mouse are the same as human rates, but use cell counts as related to humans.

Parker et al. [50] calculated δ, their model's cell proliferation rate. Their model formulation required both the cell proliferation rate and cell transfer rate to the villus, (corresponding to r_C and r_1, respectively, in our model) to be the same in order for the number of crypt cells to remain at constant equilibrium in the absence of an immune response. Their fitted δ was 0.0760 ± 0.0147 hour^{-1} ($0.00010 - 0.0015$ min^{-1}). Similarly, Barthel [7] found that the crypt-based columnar cell to transit-amplifying cell differentiation rate and crypt-based columnar cell proliferation rate must be equal for the same reason at a rate of 24.025 ± 0.006 day^{-1} (0.017 min^{-1}). However, we assumed in Section 2.5.3 that $r_1 < r_C$. We use the values from Parker et al. [50] and Barthel [7] for our r_1, cell migration rate from crypt to base of villus, and find different rates for r_C, healthy cell proliferation rate in the crypt, since both of these rates were calculated based primarily on transfer and then set equal to the rate of proliferation in the crypt in order to have a steady cell population. Additionally, we set the lower bound for r_1, the cell differentiation rate from crypt to villus base, as the lower bound from Parker et al. [50], while the upper bound comes from Barthel [7], providing the range $0.0010 - 0.0017$ min^{-1}. We assume that the cell migration rate from the crypt to the base of the villus is the same for both a healthy subject and patient with CeD.

Dickson et al. [22] found that the migratory rate of epithelial cells from the crypt to the villus surface is reduced from a range of three to five days ($0.0021 - 0.0035$ min^{-1}) in a healthy subject to one to two days ($0.00069 - 0.0014$ min^{-1}) in a patient with CeD. We use these values to calculate r_2, the cell migration rate from the base to the tip of the villus.

Barthel [7] fitted the loss rate of enterocytes, their λ_1, as 0.35364 ± 0.00007 day^{-1} (0.00025 min^{-1}). Demin et al. [21] fitted *in vitro* data to find the death rate for both activated and non-activated intestinal epithelial cells, their k_d^{iec}, as $0.015468 - 0.015474$ hour^{-1} (0.00026 min^{-1}). We use these two values to determine the range for d_V, the natural villus cell turnover rate. Since this is the natural villus cell turnover rate, it should remain the same between a healthy subject and patient with CeD.

Parker et al. [50] found that the crypt cell production rate was 1.51 ± 0.325 cells per hour ($0.020 - 0.031$ min^{-1}) in the duodenum. Using this information, we were able to determine a range of values for r_C, the healthy cell proliferation rate in the crypt. Since this is the healthy cell proliferation rate in the crypt, it should remain the same between a healthy subject and patient with CeD.

Demin et al. [21] found the death rate of activated and non-activated IELs, their k_d^{iel}, to be $0.05452 - 0.05533$ hour^{-1} ($0.00091 - 0.00092$ min^{-1}) by fitting against *in vitro* data. We use this to find γ, the natural decay rate of IELs. This will only be used in simulations for a patient with CeD since there is no immune response in a healthy subject.

According to Crosnier et al. [20], each crypt has about 250 cells while each villus about 3,500 cells. However, these values vary depending upon their position along the small bowel. When using these reference values, Barthel [7] found in the long-term, villus cells tended towards $2,844 \pm 72$ cells for a patient with CeD. Generally, a patient with CeD will have a shorter villus height and an increased crypt depth, corresponding to fewer villus cells and more crypt cells in comparison with a healthy subject. Using the long-term dynamical range from [7], we conclude that there are approximately $2,772 - 2,916$ cells in each villus for a patient with CeD. Since the standard error from [7] was approximately 2.5%, we vary the literature value given by [20] for a healthy subject by the same amount and assume there are $3,411 - 3,589$ cells in each villus for a healthy patient. Similarly, we assume there will be approximately $244 - 256$ cells in each crypt for a healthy subject. We follow the assumptions from [7] and assume that there is a maximal 30% increase in cells in the crypt for a patient suffering from CeD and assume that there will be a range between $244 - 325$ cells in each crypt for patients suffering from CeD. We can then set the range for K_C, the healthy carrying capacity of crypt cells, to be $244 - 256$ for a healthy subject and $244 - 325$ for a patient

with CeD. However, since we have split the villus cells into two compartments, more work will need to be done to calculate ranges for the transition thresholds, K_1 and K_2.

In Section 2.5.3, we found that when the immune response is absent, corresponding to a healthy subject, there is a stable equilibrium point at $(\frac{r_1 K_C}{r_2}(1-\frac{r_1}{r_C}), \frac{r_1 K_C}{d_V}(1-\frac{r_1}{r_C}), K_C(1-\frac{r_1}{r_C}))$. Thus, in the case of a healthy subject, if the initial conditions are set to this equilibrium, we will remain at this equilibrium. Thus, for a healthy subject, we set $V_{1_0} = \frac{r_1 K_C}{r_2}(1-\frac{r_1}{r_C})$ and $V_{2_0} = \frac{r_1 K_C}{d_V}(1-\frac{r_1}{r_C})$. Also in Section 2.5.3, we found that we must require three conditions:

1. $r_1 < r_C$, which holds since $\max(r_1) = 0.0017 < 0.020 = \min(r_C)$,

2. $\frac{r_1 K_C}{r_2}(1-\frac{r_1}{r_C}) < K_1$, which holds for a healthy subject when
   $\frac{\max(r_1)\max(K_C)}{\min(r_2)}(1-\frac{\min(r_1)}{\max(r_C)}) = \frac{0.0017(325)}{0.00069}(1-\frac{0.0010}{0.031}) \approx 775 < \min(K_1)$, and

3. $\frac{r_1 K_C}{d_V}(1-\frac{r_1}{r_C}) < K_2$, which holds for a healthy subject when
   $\frac{\max(r_1)\max(K_C)}{\min(d_V)}(1-\frac{\min(r_1)}{\max(r_C)}) = \frac{0.0017(325)}{0.00025}(1-\frac{0.0010}{0.031}) \approx 2,139 < \min(K_2)$.

Thus, we have that the minimum number of total villus cells in a healthy subject is 2,914. Because we want K_1 and K_2 to correspond to V_{1_0} and V_{2_0}, respectively, using our other parameter values for V_{1_0} and V_{2_0} in the equilibrium point, we find that there are approximately 27% of V_1 cells (775/2,914) and 73% of V_2 cells (2,139/2,914). Since we know the range for the total number of villus cells for both healthy subjects $(3,411 - 3,589)$ and patients with CeD $(2,772 - 2,916)$, we can calculate K_1 and K_2 using this same percentage split. Thus, we find K_1 should range between $921 - 969$ cells for a healthy subject and $748 - 787$ cells for a patient with CeD. K_2 should range between $2,490 - 2,620$ cells for a healthy subject and $2,024 - 2,129$ cells for a patient with CeD.

Next, we would like to calculate length conversion factors α and β. Adelman et al. [1] provided measurements for mucosal sections cut perpendicular to the luminal surfaces from a patient with well-treated CeD on a GFD and from two patients with incompletely treated CeD and varying degrees of architectural abnormality. Recall that a patient with well-treated CeD on a GFD is comparable to a healthy subject. The villus height provided for this patient is 524.79 μm and the crypt depth was 150.02 μm. From our ranges earlier, we found that there were between $3,411 - 3,589$ cells in the villus and $244 - 256$ cells in the crypt for

a healthy subject, averaging $3,500$ and 250 cells, respectively. Using this information and the average from the number of cells in the villus and crypt for a healthy subject, we can calculate α (524.79 μm/$3,500$ cells = 0.15 μm/cells) and β (150.02 μm/250 cells = 0.60 μm/cells), respectively. Additional calculations are completed in Section 2.6.2 to get a more accurate range for these conversion factors. We assume that the density of cells in the villus and crypt are the same between a healthy subject and patient with CeD and will thus have the same length conversion factors as their healthy counterparts. Values are provided in Table 2.5.

There remain three more parameters in the system of ODEs whose values were derived using simulations: d_1 and d_2, the destruction rate of villus cells in the base and tip by IELs, respectively, and p, the additional proliferation rate of crypt cells in the due to IELs. All three of these values can not be found in the literature and were instead calculated using model simulations. Information on how these three values were determined is described in Section 2.6.2.

2.6.2 Model Simulation Results

We use the ode45 function in MATLAB [45] to obtain numerical solutions to system of ordinary differential equations (2.1) using the average of the ranges for the parameter values presented in Table 2.4 and initial conditions as determined from the non-trivial, biologically feasible equilibrium point in the steady state analysis in Section 2.5.3.

A Heaviside function is used to turn off cell transfer when villus cellular populations reaches their respective transition threshold. In our model, we relate the number of villus cells, V_i, to their respective transition threshold, K_i, by

$$H\left(1-\frac{V_i}{K_i}\right)=\begin{cases}1 & \text{if } V_i < K_i\\ 0 & \text{if } V_i \geq K_i\end{cases} \tag{2.24}$$

for $i = 1, 2$. That is, the transfer turns off when the villus cells are above their respective transition threshold.

Table 2.5: Length conversion factors and their units for CeD model.

Parameter	**Description [Units]**	**Healthy Subject**	**Celiac Patient**	**Source**
α	length conversion factor for villus [μm/cells]	$0.15 - 0.38$	$0.15 - 0.38$	[1]
β	length conversion factor for crypt [μm/cells]	$0.59 - 0.63$	$0.59 - 0.63$	[1]

Simulations of a Healthy Subject

We begin by simulating a healthy subject to ensure that our model can represent the "baseline" scenario. We refer to a healthy subject as an individual who either does not react adversely to gluten or a patient with well-treated CeD on a GFD. Generally, a patient with well-treated CeD will have small intestinal structure comparable to a healthy subject; the only noticeable change includes an increased number of IELs [1]. However, since the IELs in our model are activated IELs, the simulations for a healthy subject and patient with well-treated CeD will be identical.

To reproduce the intestinal structure of a healthy patient with no immune response to e.g. gluten, the model will not have activated IELs or any IEL activation. Mathematically, this is represented as $I(t) \equiv 0$ and $A(t) \equiv 0$. Recall from Section 2.5.3, that our initial conditions for a healthy subject should be calculated based on our choice of other parameter values. Using the set of baseline parameters calculated as the average of the ranges given in Table 2.4, we have initial conditions $V_{1_0} = 114.15$ cells, $V_{2_0} = 1253.46$ cells, $C_0 = 236.76$ cells, and $I_0 = 0$ IELs. Figure 2.4 depicts the simulations of the model for a healthy subject for a one week period.

Next, we would like to compare these cellular counts to the diagnostic measurements as used in the Q-MARSH scale described in Section 2.3.6. In Adelman et al. [1], measurements, including the Vh, Cd, and Vh:Cd, were provided for four representative patients. The first, a patient with well-treated CeD on a GFD, can be compared with a healthy subject. We use the initial conditions provided above and these measurements (Vh = 524.79 μm, Cd = 150.02 μm, and Vh:Cd = 3.50) to calculate conversion constants $\alpha = 0.38$ and $\beta = 0.63$ which relate the cell counts for the villus and crypt to the Vh and Cd, respectively. Using these conversion constants, we plot the diagnostic measurements over time in Figure 2.5 for a one week period.

In Figures 2.4 and 2.5, there are minor fluctuations in each of the classes, which corresponds to the natural cellular turnover in the gut. The cellular populations and diagnostic measurements remain close to a steady state. Together, these baseline simulations produce appropriate gut dynamics of a healthy individual.

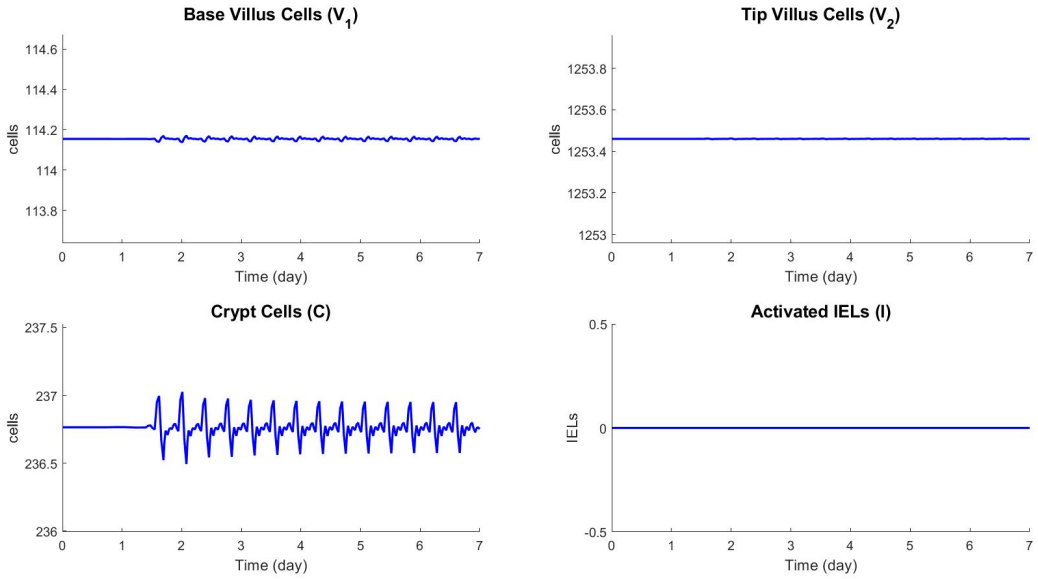

Figure 2.4: Model Simulations of Cellular Counts for a Healthy Subject. The panels display (clockwise from top left) the model variables V_1, V_2, IELs, and C over time for a healthy subject, where $I(t) = 0$ and $A(t) = 0$ in the model for this scenario to emulate the absence of immune activation. Initial conditions used for the simulation are $V_{1_0} = 114.15$ cells, $V_{2_0} = 1253.46$ cells, $C_0 = 236.76$ cells, and $I_0 = 0$ IELs. Parameter values are chosen to be the mean of the ranges given in Table 2.4.

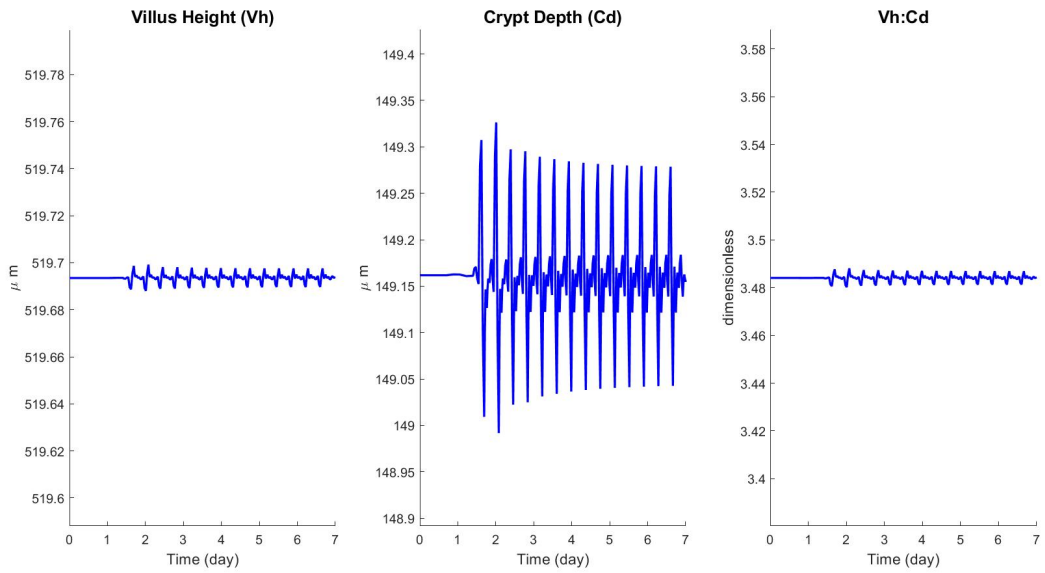

Figure 2.5: Simulation-Based Diagnostic Measurement Values for a Healthy Subject. Conversion constants $\alpha = 0.38$ and $\beta = 0.63$ were used to relate cellular densities of V_1, V_2, and C in Figure 2.4 to villus height (Vh, first panel) and crypt depth (Cd, second panel). The ratio Vh:Cd in the third panel provides the simulation result of the diagnostic histologic measurement.

Simulations with Immune Activation

It is known that among patients with CeD, each react differently to gluten in both symptoms and histologic changes. While gluten elicits an immune response, the upstream events leading to this response are not well-known and there are other causes of gut inflammation besides gluten exposure. Thus, our approach is to model scenarios of immune activation that could be due to gluten exposure or some other upstream stimulus. In general, we assume that the stimulus activates IELs, leading to mucosal damage as discussed earlier. We define an activation function for this process that will govern the source of IELs in the model. Since the upstream events regarding sources and dynamics of activation are not well-known, we consider various forms of activation. We will use the following general form for the activation rate function

$$A(t) = a_1 e^{\frac{-(t-a_2)^2}{a_3}} \tag{2.25}$$

where a_1, a_2, and a_3 are tuning constants.

The structure of $A(t)$ in (2.25) allows us to tune three function constants to consider in the activation of IELs - the magnitude of activation a_1, the time of peak activation a_2, and the duration of non-trivial activation a_3. Tuning constant a_1 determines the magnitude, or the height of the peak, of the curve. We can interpret this as the maximum impact in the activation of IELs. Tuning constant a_2 determines the time at which the activation function will reach its peak. Finally, tuning constant a_3 determines the duration of the non-trivial activation. Larger values of a_3 will result in a slower decay in activation and a longer duration of non-trivial IEL activation. Figure 2.6 depicts $A(t)$ with varying tuning constants.

To measure the effect of IEL activation in a model simulation, we calculate the area under the curve of $A(t)$. The benefit of an activation that can be decoupled from a specific stimulus (e.g., gluten) allows this work to be extended to other diseases deemed "leaky gut." However, there is some data from the literature that we later try to incorporate into our model simulations regarding IEL levels in patients with CeD on a GFD who were challenged with daily gluten exposures.

The initial conditions for a patient with CeD are chosen according to the method described earlier for a healthy subject. In the absence of gluten, a patient with CeD

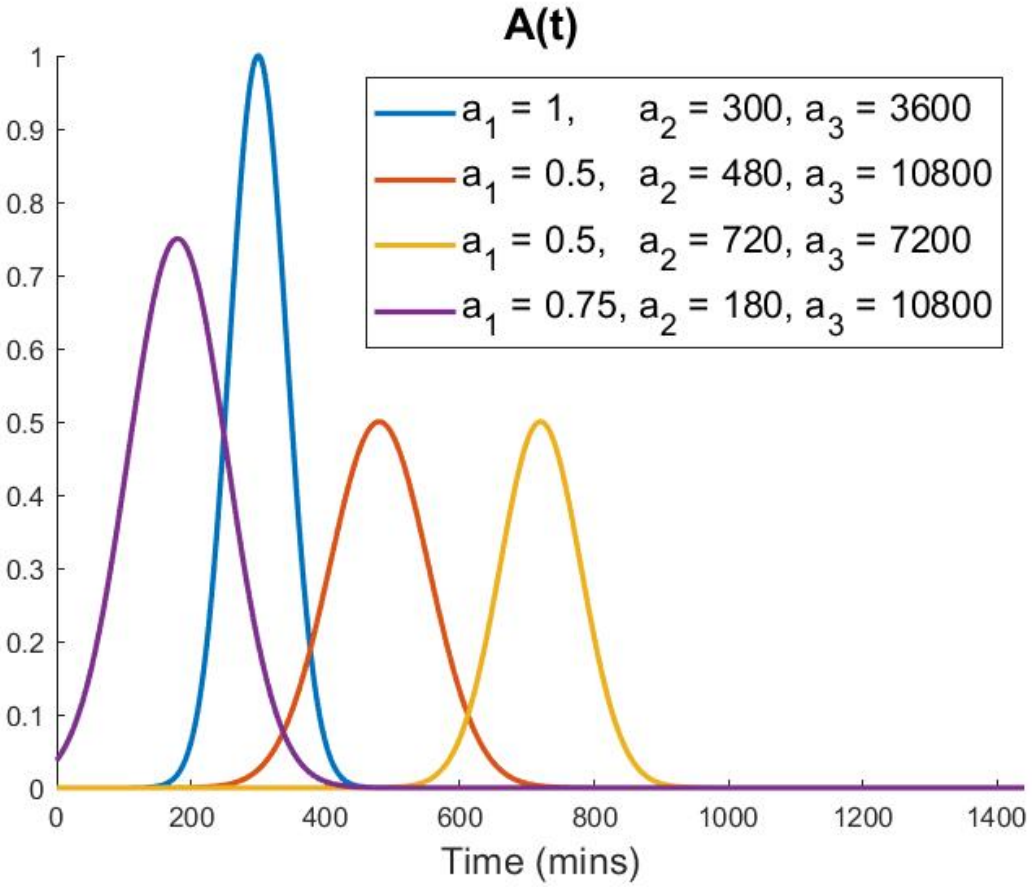

Figure 2.6: Activation function $A(t)$ used as a source term for the intraepithelial lymphocytes (IELs) in the model. Several examples of $A(t)$ are shown here with various values for tuning constants as given in the legend.

should have dynamics similar to a healthy subject, with minor fluctuations in the cellular populations and a stable Vh:Cd ratio value. Thus, when using the set of baseline parameters for a patient with CeD, defined to be the average of the ranges given in Table 2.4, we have that $V_{1_0} = 348.08$ cells, $V_{2_0} = 1426.44$ cells, $C_0 = 269.44$ cells, and $I_0 = 0$ IELs.

We have already simulated the response for a healthy subject; that is, without IEL dynamics. Now, when including IEL dynamics, we incorporate terms and associated parameters for the destruction of the villus cells in the base and tip by IELs, with rates d_1 and d_2, respectively. In addition, we include an additional proliferation of crypt cells in the presence of IELs with growth rate, p. Initially, we choose $d_1 = 0$ and vary the other two parameters. The literature notes that the damage done to the cells in the base of the villus is negligible, leading us to set $d_1 = 0$. The parameters d_2 and p cannot be found in the literature as not much is known exactly how the immune response causes villi damage and crypt overproliferation in a patient suffering from CeD, nor are there known rates at which this damage occurs. We can explore varying rates to determine how disease progression and recovery are affected when the immune response is varied.

By varying IEL activation through activation function $A(t)$ and the immune response through parameters d_2 and p, we investigate model scenarios for patients with incompletely-treated CeD. In these initial simulations, we simulate a single, transient activation of the IELs to observe the length of time it takes in the simulation for the gut to heal. In these simulations, we consider varying the tuning constants in the activation function for an early, delayed, and damped activation as well as varied parameter values for d_2 and p in the immune response, both weak and strong.

For the early immune activation, we will use tuning constants $a_1 = 1$, $a_2 = 180$, and $a_3 = 7,250$ in $A(t)$ as shown in solid blue in Figure 2.7. For the delayed immune activation, we use tuning constants $a_1 = 1$, $a_2 = 720$, and $a_3 = 7,250$ as shown in dotted red. For the damped immune activation, we use tuning constants $a_1 = 0.5$, $a_2 = 720$, and $a_3 = 29,000$ as shown in dashed green. All three of these activation functions have the same area under the curve for the first 24 hours (150.919). Thus, each will cause the same increase number of IELs over the 24 hour period. However, what will vary is the rate at which this activation

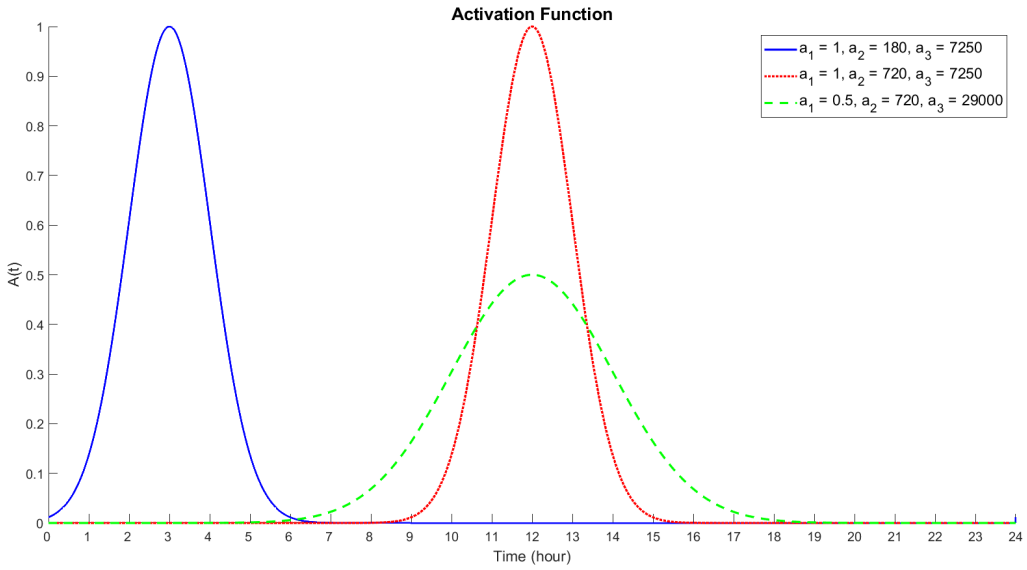

Figure 2.7: Three considered IEL activation functions simulated over a one day period. All three functions have the same area under the curve for the first 24 hours. Solid blue corresponds to an early immune activation, dotted red corresponds to a delayed immune activation, and dashed green corresponds to a damped immune activation.

happens over time. Coupled with γ, the natural decay rate of IELs, the IELs may not be at the same level for all time when considering these three activation functions.

In these simulations, we will consider "extreme" immune responses to begin to hone in on what values make sense for the parameters d_2, the additional destruction of villus cells in the tip by IELs, and p, the additional proliferation of crypt cells due to IELs. We will relate both of these parameters to the respective growth rates of the compartments, r_2 and r_C. For a weak immune response, we set d_2 to be 10% of r_2 and p to be 10% of r_C. For a strong immune response, we set d_2 to be 110% of r_2 and p to be 110% of r_C. By considering these extremes, we hope to narrow a feasible range for these parameter values.

We begin by considering a patient with an early activation and weak immune response. Figure 2.8 represents the cellular dynamics after a single, transient activation on the first day and Figure 2.9 depicts the changes in the diagnostic features for this patient. The red horizontal dashed lines correspond to the time course of a healthy subject with the same initial conditions. The multicolored trajectory represents the patient with CeD with change in color corresponding to change in day.

While we might expect that the cells in the base of the villus decrease due to villus atrophy, instead the cells in the base of the villus increase fairly quickly in the first day and begin to level off, remain steady through day two, and begin their decline on day three. These cells begin to reach their healthy level around day seven. Conversely, the cells in the tip of the villus decay drastically in the first day and begin a steady increase over the next two weeks, finally reaching near healthy at the conclusion of day fourteen. The crypt cells are able to rebound more quickly and are back to a healthy state after their quick increase by the conclusion of day four. The IELs behave similarly to the crypts.

Using the same conversion constants from the healthy subject simulations, $\alpha = 0.38$ and $\beta = 0.63$, we can relate the cellular turnover to the Vh:Cd for this simulated patient. In Figure 2.9, the Vh:Cd declines from 3.97 to 0.14 in the first day, but in the first day begins to approach its healthy state of 3.97. By the conclusion of day 18, the gut structure has been restored. Thus, simulations depict that for this patient with CeD, after a single, transient, and early immune activation with a weak immune response, it would take approximately 18 days for the gut to be restored.

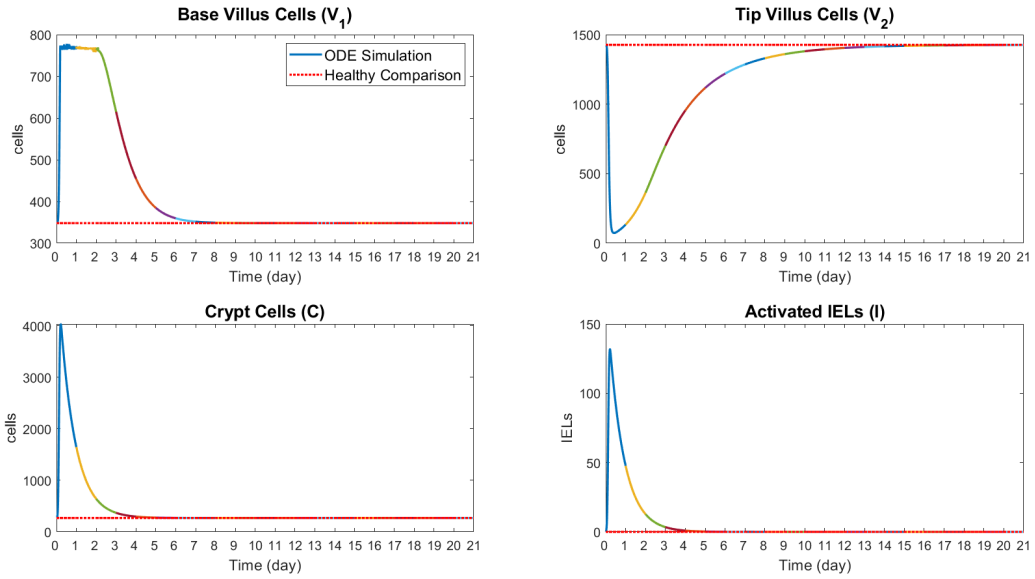

Figure 2.8: Model Simulation of Cellular Counts for a Patient with CeD - Single, Early Activation and Weak Immune Response. The panels display (clockwise from top left) the model variables V_1, V_2, IELs, and C over time for a patient with CeD. Initial conditions used for the simulation are $V_{1_0} = 348.08$ cells, $V_{2_0} = 1426.44$ cells, $C_0 = 269.44$ cells, and $I_0 = 0$ IELs. Parameter values are chosen to be the mean of the ranges given in Table 2.4. For the immune response, $A(t)$ has tuning constants $a_1 = 1$, $a_2 = 180$, and $a_3 = 7,250$ as shown in Figure 2.7, $d_2 = 0.1r_2$, and $p = 0.1r_C$.

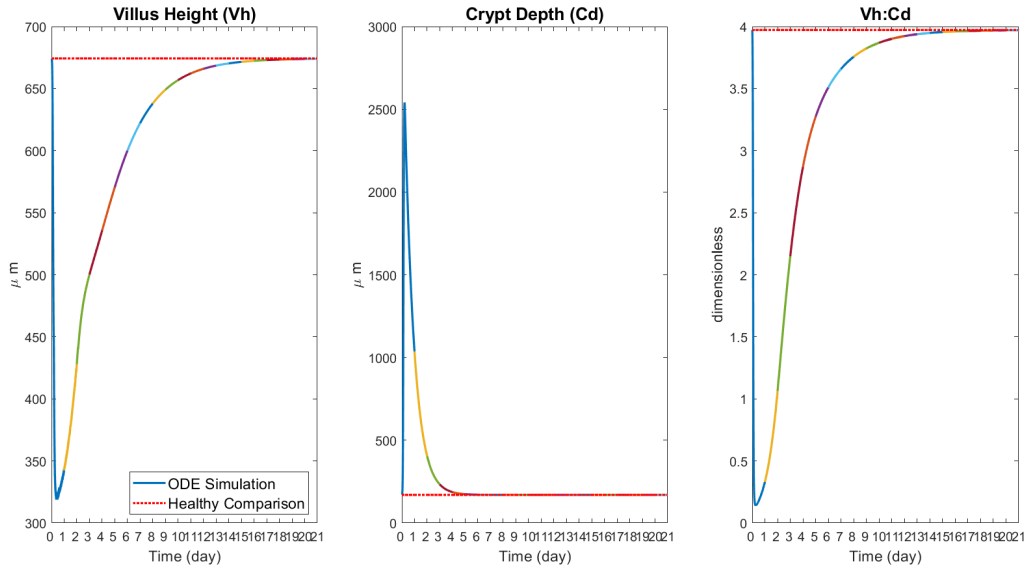

Figure 2.9: Simulation-Based Diagnostic Measurement Values for a Patient with CeD - Single, Early Activation and Weak Immune Response. Conversion constants $\alpha = 0.38$ and $\beta = 0.63$ were used to relate cellular densities of V_1, V_2, and C in Figure 2.8 to provide the simulation results of the diagnostic measurements, Vh, Cd, and Vh:Cd.

In Section 2.3.4, it was mentioned that the amount of time that it takes the villi to heal is highly variable and can take anywhere from one to two weeks to several years, while complete healing never occurs for some individuals. Thus, the parameters that we utilize in our mathematical model depict a relatively fast healing rate of the small intestine and it should be noted our results may apply to a limited set of individuals. The parameters involved in this healing process were calculated based upon parameters used in other models involved in cellular turnover. With cellular count data from patients with CeD, these rates could be calculated for a larger variety of patients and our model could then depict a slower healing rate where it takes longer than 18 days for the the gut to be restored.

We now consider a strong immune response with the same single, transient, early activation function. We set d_2 to be 1.1 times r_2 and p to be 1.1 times of r_C. Figure 2.10 shows the time courses for this simulated patient. In considering these dynamics, we notice that the cells in the base and tip of the villus follow similar trajectories to those in Figure 2.8 for the weak immune response. However, we see that the number of crypt cells increases to a maximum of 41,534.41 cells which is not biologically feasible. Thus, we will no longer consider this strong immune response.

When we consider a single, transient, delayed activation with a weak immune response, we see the same results as the earlier activation. This should not be unexpected as we are using the same activation function, but with a nine hour delay. Figures 2.11 and 2.12 depict the time courses when the peak activation is delayed until twelve hours from the start of each day. Thus, it does not appear that when the activation function $A(t)$ reaches its peak time of activation affects our results.

Now, we simulate the damped response. In this case, we maintain the delay where the time of peak activation happens 12 hours into the day, but the magnitude of activation is half of what it was before, meaning at the duration of non-trivial activation must be increased so that there is the same area under the curve, corresponding to the same number of activated IELs each day. In Figures 2.13 and 2.14, we see the time courses when the peak activation is delayed until twelve hours from the start of the first day, but with the damped activation.

In comparing these simulations to the delayed response, we note that the villus cell dynamics do not alter. However, we see a reduced peak in the number of crypt cells as well

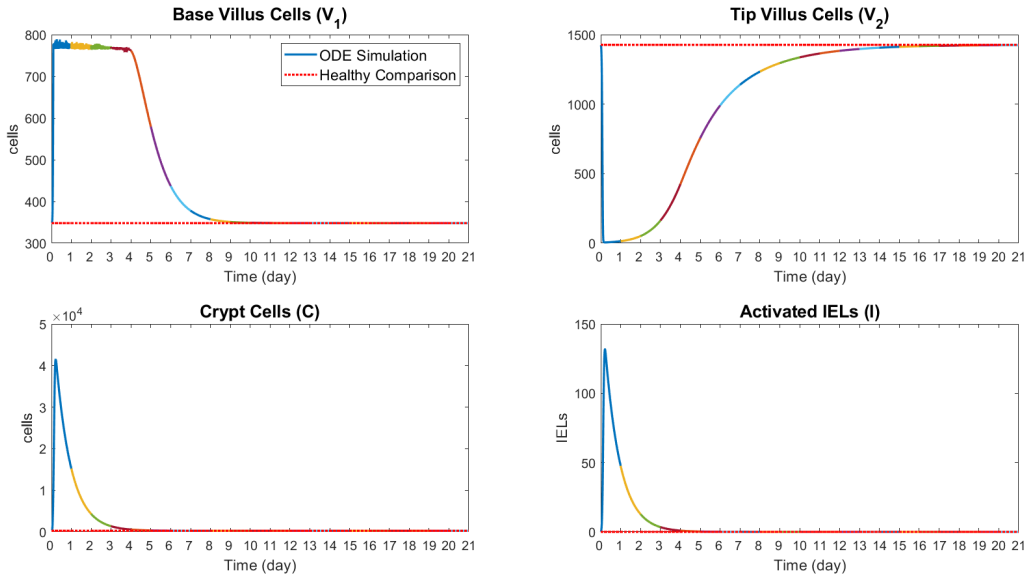

Figure 2.10: Model Simulation of Cellular Counts for a Patient with CeD - Single, Early Activation and Strong Immune Response. The panels display (clockwise from top left) the model variables V_1, V_2, IELs, and C over time for a patient with CeD. Initial conditions used for the simulation are $V_{1_0} = 348.08$ cells, $V_{2_0} = 1426.44$ cells, $C_0 = 269.44$ cells, and $I_0 = 0$ IELs. Parameter values are chosen to be the mean of the ranges given in Table 2.4. For the immune response, $A(t)$ has tuning constants $a_1 = 1$, $a_2 = 180$, and $a_3 = 7,250$ as shown in Figure 2.7, $d_2 = 1.1r_2$, and $p = 1.1r_C$.

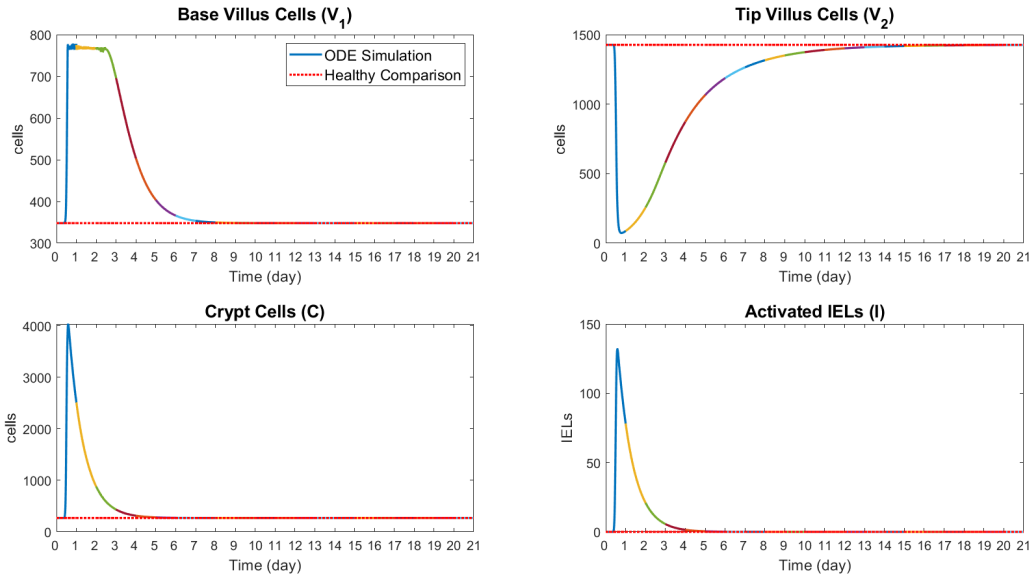

Figure 2.11: Model Simulation of Cellular Counts for a Patient with CeD - Single, Delayed Activation and Weak Immune Response. The panels display (clockwise from top left) the model variables V_1, V_2, IELs, and C over time for a patient with CeD. Initial conditions used for the simulation are $V_{1_0} = 348.08$ cells, $V_{2_0} = 1426.44$ cells, $C_0 = 269.44$ cells, and $I_0 = 0$ IELs. Parameter values are chosen to be the mean of the ranges given in Table 2.4. For the immune response, $A(t)$ has tuning constants $a_1 = 1$, $a_2 = 720$, and $a_3 = 7,250$ as shown in Figure 2.7, $d_2 = 0.1r_2$, and $p = 0.1r_C$.

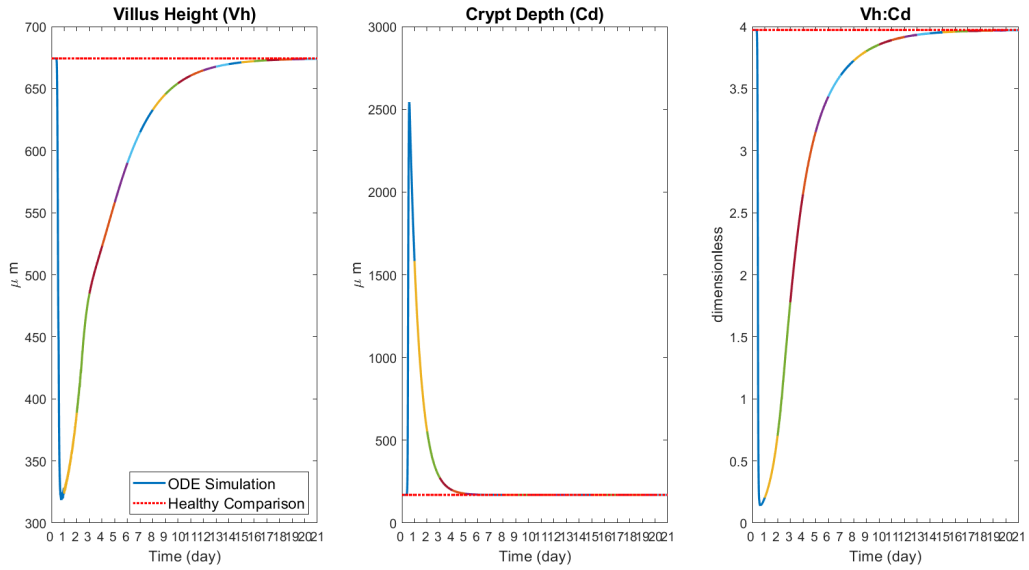

Figure 2.12: Simulation-Based Diagnostic Measurement Values for a Patient with CeD - Single, Delayed Activation and Weak Immune Response. Conversion constants $\alpha = 0.38$ and $\beta = 0.63$ were used to relate cellular densities of V_1, V_2, and C in Figure 2.11 to provide the simulation results of the diagnostic measurement, Vh:Cd.

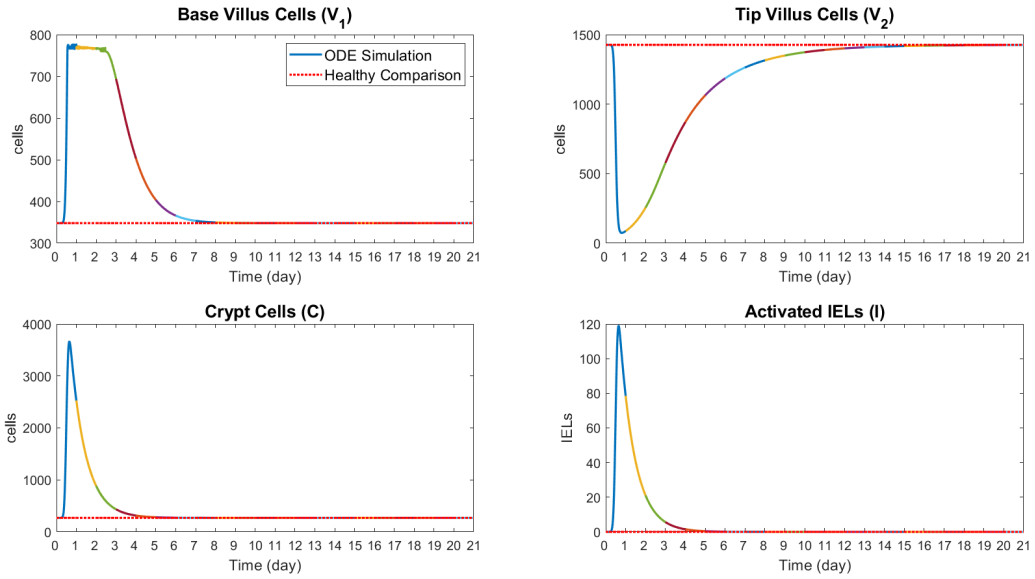

Figure 2.13: Model Simulation of Cellular Counts for a Patient with CeD - Single, Damped Activation and Weak Immune Response. The panels display (clockwise from top left) the model variables V_1, V_2, IELs, and C over time for a patient with CeD. Initial conditions used for the simulation are $V_{1_0} = 348.08$ cells, $V_{2_0} = 1426.44$ cells, $C_0 = 269.44$ cells, and $I_0 = 0$ IELs. Parameter values are chosen to be the mean of the ranges given in Table 2.4. For the immune response, $A(t)$ has tuning constants $a_1 = 0.5$, $a_2 = 720$, and $a_3 = 29,000$ as shown in Figure 2.7, $d_2 = 0.1r_2$, and $p = 0.1r_C$.

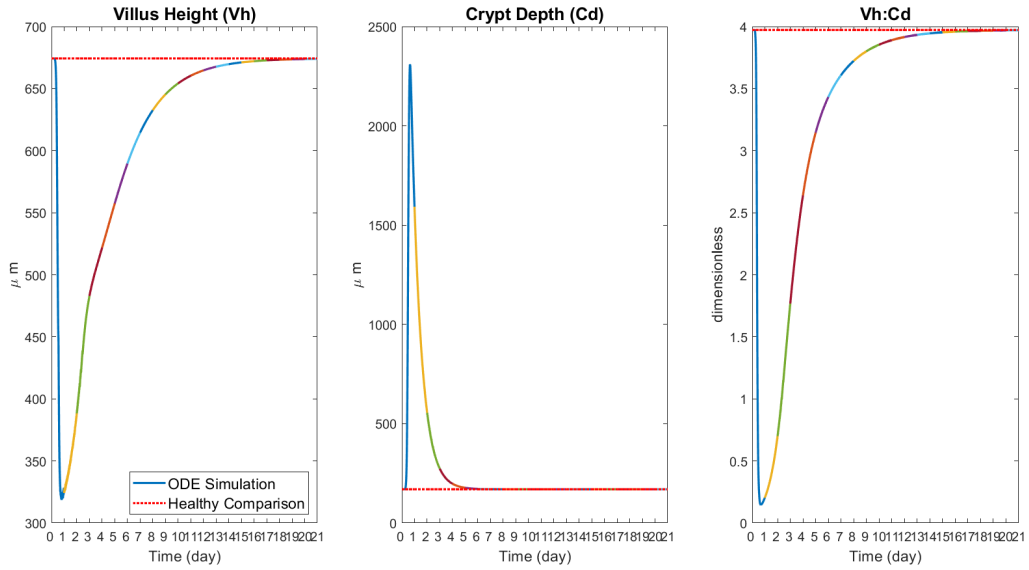

Figure 2.14: Simulation-Based Diagnostic Measurement Values for a Patient with CeD - Single, Damped Activation and Weak Immune Response. Conversion constants $\alpha = 0.38$ and $\beta = 0.63$ were used to relate cellular densities of V_1, V_2, and C in Figure 2.13 to provide the simulation results of the diagnostic measurement, Vh:Cd.

as IELs. In the delayed response, we see a maximum of 4,037 crypt cells and 132 IELs, whereas in the damped response, we see a maximum of 3,667 crypt cells and 119 IELs. Due to the dampening of the crypt cells, we see a reduction in the crypt depth, but not enough that the Vh:Cd is effected. Thus, when we lessen the magnitude of activation, the natural decay of the IELs has the ability to reduce the IELs and thus reduce the downstream effects.

Next, we simulate a simulated patient with CeD undergoing a gluten challenge in which the patient ingests gluten every day at the same time each day. These challenges are typical in clinical trials to aid in determining whether a potential therapeutic is viable in treating CeD. In order to mathematically model this daily activation, we pulse the activation functions shown in Figure 2.7 every 1,440 minutes. Figure 2.15 represents the first same activation function for the IELs (solid blue) pulsed over a seven day period, representing activation of IELs due to a seven day gluten challenge.

We will consider a patient with only a weak immune response since the strong immune response in the single, transient activation did not make sense biologically. Recall, we considered only extreme immune responses to determine what values of d_2 and p make sense. Now, we only consider values closer to the lower end of this range. However, we will still consider the effects of the early, delayed, and damped immune activation. Figure 2.16 depicts the cellular dynamics of a patient with an early activation and weak immune response involved in a gluten challenge. Over the seven day time course, the cells in the base of the villus increase during the first day and then hover around 769.45 cells. The cells in the tip of the villus decline on the first day with an attempt to restore to a healthy level, but with each subsequent activation of the IELs, continue to decline. The crypt cells increase every day due to the activation of the IELs with an attempt to recover to a healthy state before the next activation causes another increase in the amount of cells. Finally, the activated IELs increase the most in the first half of the day as the activation function $A(t)$ reaches its peak three hours after the start of each day.

Relating these cellular dynamics to the diagnostic procedures, we can consider the time course for the Vh:Cd for this patient. In Figure 2.17, we see that following the first activation of the IELs, the Vh:Cd takes the greatest decline from 3.97, corresponding to a patient with CeD adhering to a GFD, to 0.25, a ratio corresponding to an unhealthy patient. Each day,

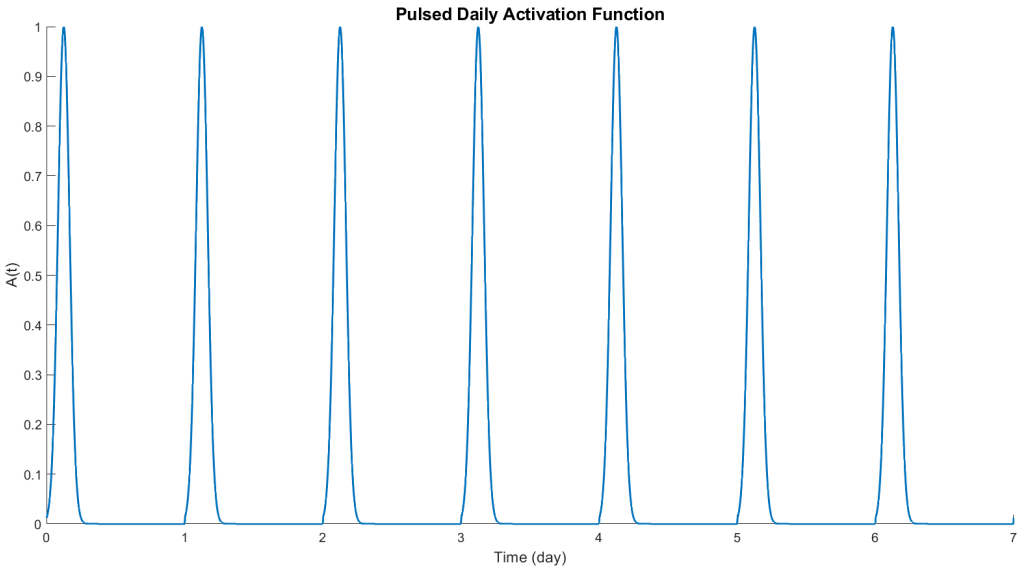

Figure 2.15: Potential IEL activation function pulsed daily to represent a seven day gluten challenge. Activation function $A(t)$ corresponds to the solid blue function in Figure 2.7. Tuning constants used include $a_1 = 1$, $a_2 = 180$, and $a_3 = 7,250$.

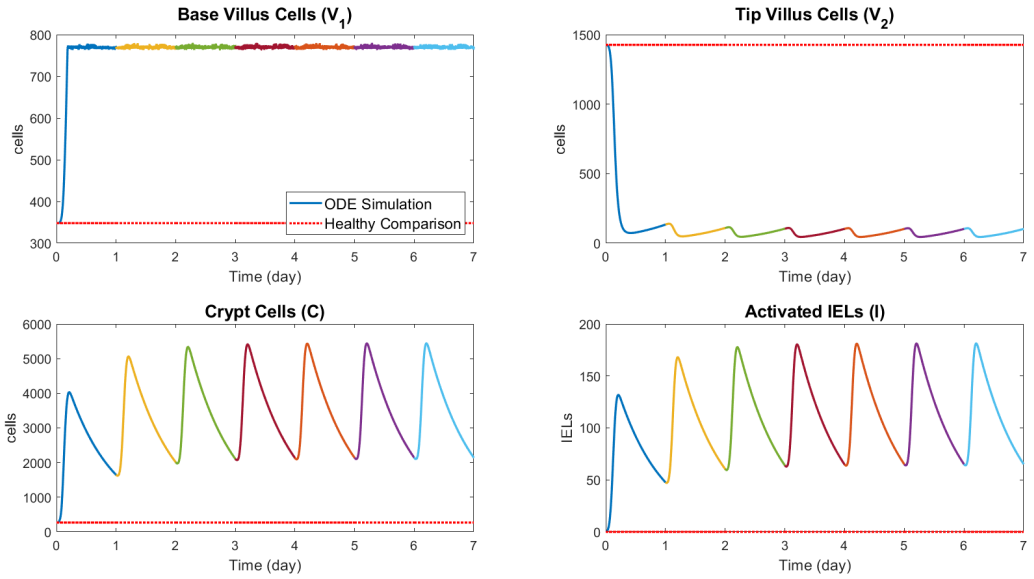

Figure 2.16: Model Simulation of Cellular Counts for a Patient with CeD - Daily, Early Activation and Weak Immune Response. The panels display (clockwise from top left) the model variables V_1, V_2, IELs, and C over time for a patient with CeD. Initial conditions used for the simulation are $V_{1_0} = 348.08$ cells, $V_{2_0} = 1426.44$ cells, $C_0 = 269.44$ cells, and $I_0 = 0$ IELs. Parameter values are chosen to be the mean of the ranges given in Table 2.4. For the immune response, $A(t)$ has tuning constants $a_1 = 1$, $a_2 = 180$, and $a_3 = 7,250$ as shown in Figure 2.7 is pulsed daily, $d_2 = 0.1r_2$, and $p = 0.1r_C$.

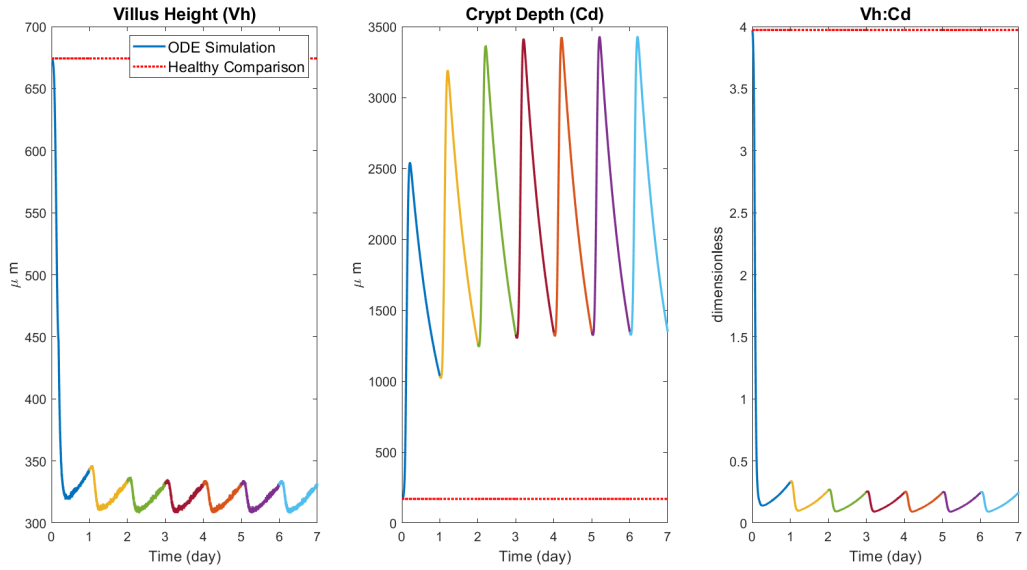

Figure 2.17: Simulation-Based Diagnostic Measurement Values for a Patient with CeD - Daily, Early Activation and Weak Immune Response. Conversion constants $\alpha = 0.38$ and $\beta = 0.63$ were used to relate cellular densities of V_1, V_2, and C in Figure 2.16 to provide the simulation results of the diagnostic measurement, Vh:Cd.

the Vh:Cd begins to move back towards its healthy state, but declines again due to the next activation of IELs.

Next, we will simulate this same weak immune response, but delay the time of peak activation to be 12 hours after the start of the the day. Figure 2.18 depicts the trajectories of the cellular populations. We note that the simulations are similar to those with the early immune response, but all cellular populations have a slight delay since we are again using the same activation function with a time delay. These similarities continue to Figure 2.19 in the simulations of the histologic measurements.

Finally, we consider a damped activation. In Figures 2.20 and 2.21, we see the time courses when the peak, daily activation is delayed until twelve hours from the start of each day, but with the damped activation. Here, we reduce the magnitude of the peak activation in half and lengthen the duration of non-trivial activation in order to maintain the area under $A(t)$ over the first 24 hours. In comparing these simulations to the delayed daily activation, again see very little change.

In comparing all three activation functions, we see that the dynamics of the cellular counts and thus the diagnostic measurements remain the same. This is likely because all three functions have the same area under $A(t)$ over a 24 hour period. We can conclude that the form of activation matters less than the area under the curve, especially with repeated activations.

These simulations show the harm in repeated exposure to gluten for a patient suffering from CeD. More than that, these simulations depict that the damage to the gut at the end of day two will be the same as the end of three, and so on. Typically, gluten challenges last upward of a month. It may be possible to shorten the duration of gluten challenges and still receive the same histologic results regarding damage and healing of the gut. In order to verify this claim, more data from biopsies would be needed, potentially collected weekly or bi-weekly, to determine if the Vh:Cd measurements from individual patients follow these results.

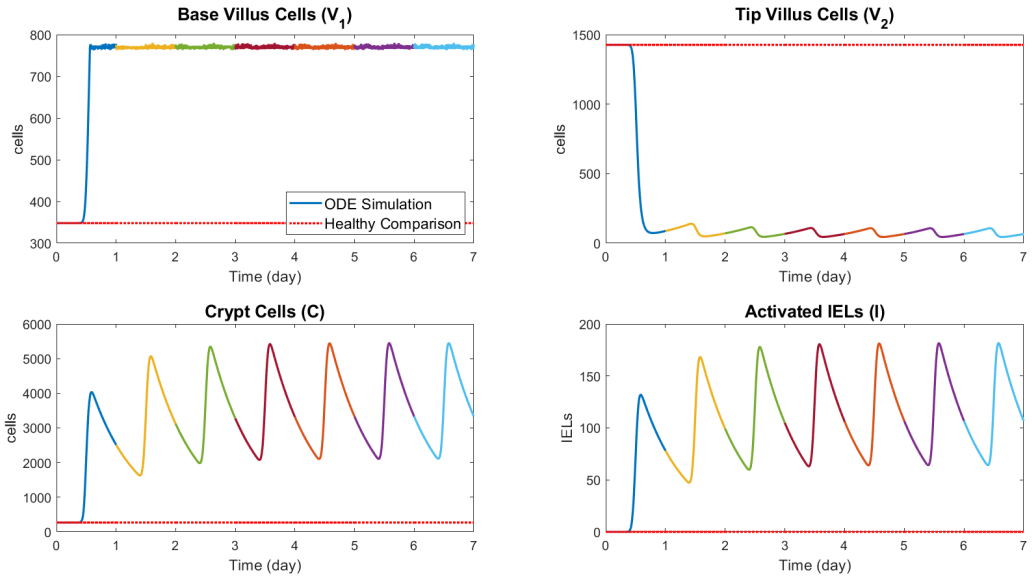

Figure 2.18: Model Simulation of Cellular Counts for a Patient with CeD - Daily, Delayed Activation and Weak Immune Response. The panels display (clockwise from top left) the model variables V_1, V_2, IELs, and C over time for a patient with CeD. Initial conditions used for the simulation are $V_{1_0} = 348.08$ cells, $V_{2_0} = 1426.44$ cells, $C_0 = 269.44$ cells, and $I_0 = 0$ IELs. Parameter values are chosen to be the mean of the ranges given in Table 2.4. For the immune response, $A(t)$ has tuning constants $a_1 = 1$, $a_2 = 720$, and $a_3 = 7,250$ as shown in Figure 2.7 is pulsed daily, $d_2 = 0.1r_2$, and $p = 0.1r_C$.

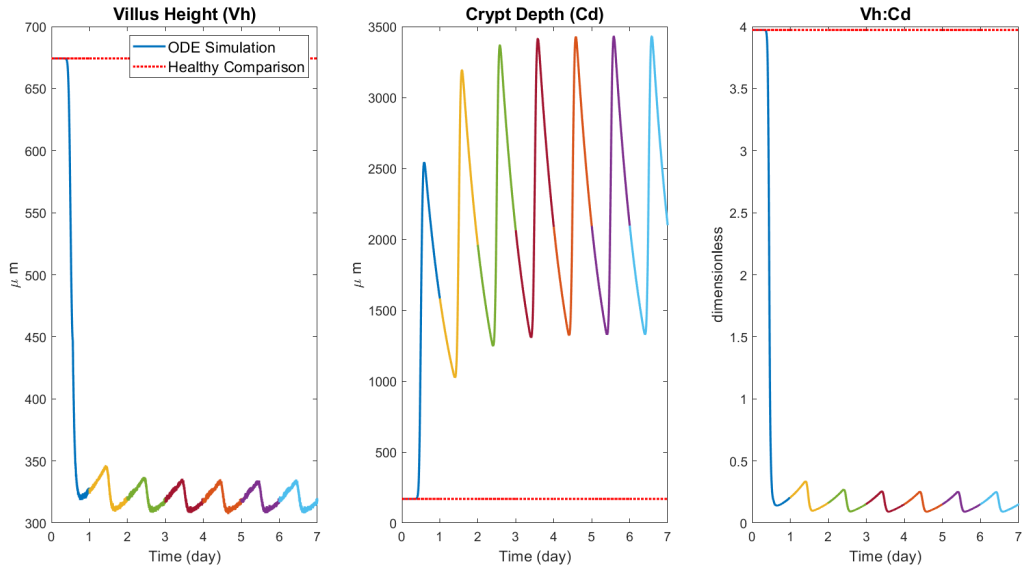

Figure 2.19: Simulation-Based Diagnostic Measurement Values for a Patient with CeD - Daily, Delayed Activation and Weak Immune Response. Conversion constants $\alpha = 0.38$ and $\beta = 0.63$ were used to relate cellular densities of V_1, V_2, and C in Figure 2.18 to provide the simulation results of the diagnostic measurement, Vh:Cd.

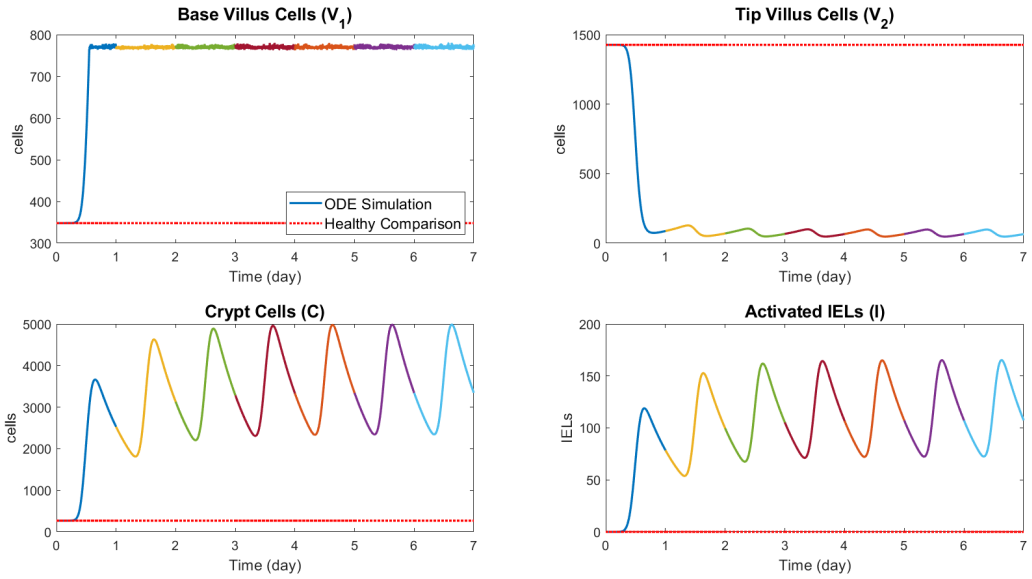

Figure 2.20: Model Simulation of Cellular Counts for a Patient with CeD - Daily, Damped Activation and Weak Immune Response. The panels display (clockwise from top left) the model variables V_1, V_2, IELs, and C over time for a patient with CeD. Initial conditions used for the simulation are $V_{1_0} = 348.08$ cells, $V_{2_0} = 1426.44$ cells, $C_0 = 269.44$ cells, and $I_0 = 0$ IELs. Parameter values are chosen to be the mean of the ranges given in Table 2.4. For the immune response, $A(t)$ has tuning constants $a_1 = 0.5$, $a_2 = 720$, and $a_3 = 29,000$ as shown in Figure 2.7 is pulsed daily, $d_2 = 0.1r_2$, and $p = 0.1r_C$.

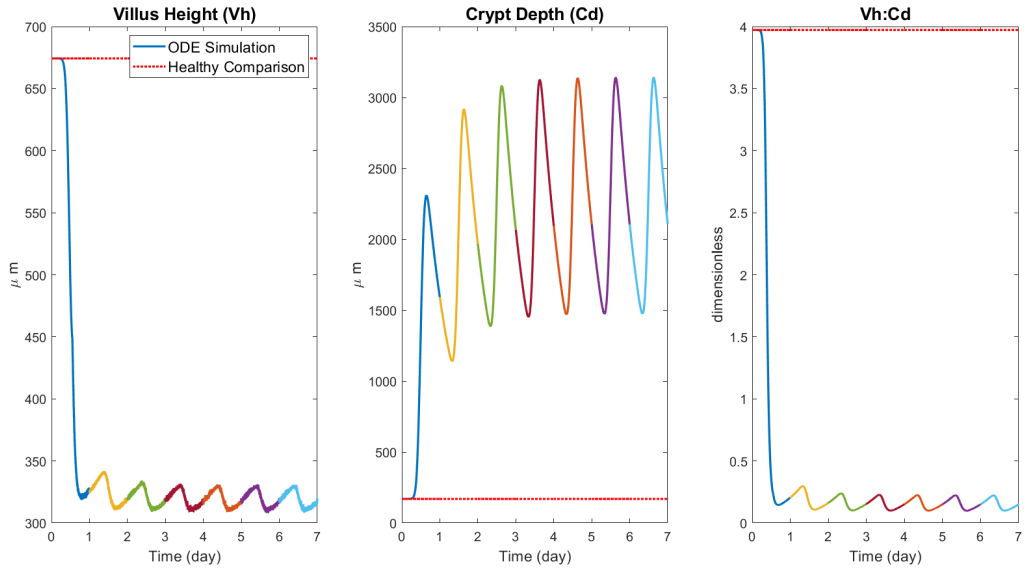

Figure 2.21: Simulation-Based Diagnostic Measurement Values for a Patient with CeD - Daily, Damped Activation and Weak Immune Response. Conversion constants $\alpha = 0.38$ and $\beta = 0.63$ were used to relate cellular densities of V_1, V_2, and C in Figure 2.20 to provide the simulation results of the diagnostic measurement, Vh:Cd.

2.6.3 Data from Gluten Challenges

Drug-development programs often include gluten challenges in which patients with diagnosed CeD ingest gluten on (typically) a daily basis to determine whether the therapeutic intervention decreases the amount of gut damage in comparison to a control group. Lähdeaho et al. [36] challenged 25 adults with treated CeD with gluten to assess the amount needed to cause small-bowel mucosal deterioration. These adults were challenged with low ($1-3$ g) or moderate ($3-5$ g) doses of gluten daily for 12 weeks. Symptoms, small-bowel morphology, densities of CD3+ IELs, and celiac serology were determined. Mucosal morphological changes were measured via the Vh:Cd and the inflammatory aspect was measured in terms of the density of mucosal IELs, both known to be sensitive, continuous functional parameters of gluten ingestion in CeD [36].

The study found that both moderate and low amounts of gluten induced small-bowel morphological damage in 67% of patients with CeD. Moderate gluten doses also triggered mucosal inflammation and more gastrointestinal symptoms leading to premature withdrawals in seven cases; low amounts of gluten also caused significant mucosal deterioration in the majority of the patients. In 22% of those who developed significant small-intestinal damage, symptoms remained absent [36].

Patients had been adhering to a strict GFD for at least two years and in clinical remission. Upper gastrointestinal endoscopy with duodenal biopsies, laboratory analysis, and clinical evaluation including gastrointestinal symptoms were carried out prior to and following the challenge. The complete challenge lasted 12 weeks (84 ± 14 days, range $29-103$ days). The study patients were divided into two groups, those consuming a moderate ($3-5$ g) and low ($1-3$ g) amount of gluten daily [36].

Each patient consumed gluten as a biscuit daily during the challenge while otherwise continuing a strict GFD. Up to seven small-bowel biopsy specimens were taken form the distal part of the duodenum upon upper gastrointestinal endoscopy. Morphometric analysis measuring Vh:Cd was made in well-oriented biopsy samples and a ratio less than 2.0 was regarded as compatible with villous atrophy and crypt hyperplasia and indicative of active CeD. During the gluten-challenge, a decrease in Vh:Cd of 0.5 or more was considered

significant and equivalent to clinical gluten sensitivity. CD3+ IELs expressed as cells/mm of epithelium, the reference values being set at 37 cells/mm for CD3+ IELs. After the gluten challenge, an over 30% increase in IEL counts was considered significant and equivalent to gluten sensitivity [36].

Of the patients that began the study, 21 patients remained until conclusion. The data for these patients is included in Table 2.6. A total of ten patients had moderate daily gluten doses; the remaining eleven patients had low daily gluten doses. The data from Lähdeaho et al. [36] is useful in that we have a large sample of representative patients with varying amounts of gluten intake with comparisons between the Vh:Cd and IEL counts at both the initial and final time. Also included in the table is the change in density of IELs, which we will later consider as the number of activated IELs at the final time, corresponding to zero activated IELs at the initial time. We will not use the data from the seven patients who did not have a clinically significant reduction in Vh:Cd, highlighted in gray.

At the baseline of the study, all patients except one, despite adherence to a strict long-term GFD, had normal Vh:Cd (median 2.9, range $1.3 - 4.2$). According to the literature, such cases exist and it has been shown that 80% of treated patients may have ongoing villous atrophy despite adherence to a strict GFD. These persistent mucosal changes may be due to constant residual gluten contamination in food, which indicates the importance of alternative treatment options [36].

A clinically significant change in the density of CD3+ IELs was seen in 14 (67%) subjects; eight in the moderate and six in the low dose group. An additional four patients have been grayed out to indicate that there was not a significant increase in the density of IELs. In the moderate gluten dose group, the IEL count increased significantly from a mean of 49 cells/mm (95% CI: $38 - 60$) to 88 cells/mm (95% CI: $70 - 106$). In the low-dose group, the increase did not reach statistical significance, from a mean 78 cells/mm (95% CI: $56 - 100$) to 98 cells/mm (95% CI: $61 - 135$). Thus, daily gluten intake did not correlate with increased inflammation. Prior to this study, it was thought that IELs accumulate in the CVEU in a dose-dependent manner [36].

There were some patients with CeD who did not respond in terms of any of the measured outcome variables during the gluten challenge. It is known that patients with CeD show

Table 2.6: Data from Lähdeaho et al. [36]. Data for each patient includes the mean daily gluten intake, duration of the gluten challenge, Vh:Cd and IEL density at both the initial and final times of the study. Gray rows indicate that there was not a significant decrease in Vh:Cd and/or significant increase in IEL density.

Patient	Gluten intake [g]	Duration [days]	Vh:Cd [dim.]		IELs [cells/mm]		
			Initial	Final	Initial	Final	Change
1	5.0	29	2.8	0.8	33	88	55
2	4.9	84	3.5	3.4	43	84	41
3	4.9	38	2.9	1.3	64	128	64
4	4.7	45	2.7	0.2	70	113	43
5	4.1	61	3.0	3.5	41	56	15
6	4.0	91	3.0	0.6	45	97	52
7	3.6	91	2.8	1.4	42	125	83
8	3.6	84	3.0	2.6	38	47	9
9	3.4	88	3.8	2.3	69	83	14
10	3.3	86	2.7	0.6	68	96	28
11	2.8	89	2.5	3.1	87	52	—
12	2.7	84	3.0	0.6	87	117	30
13	2.6	81	2.9	3.1	49	50	1
14	2.4	85	2.7	1.9	146	271	125
15	2.2	103	4.2	1.3	23	48	25
16	2.1	93	1.3	0.1	130	99	—
17	2.1	85	3.3	2.4	100	79	—
18	2.1	84	2.9	1.7	75	78	3
19	2.1	83	3.2	3.4	70	123	53
20	1.4	77	3.4	3.0	41	87	46
21	1.3	78	2.5	0.8	53	71	18

varying sensitivity to gluten and the process of mucosal deterioration may take years or even decades [36]. After removing the patients who did not have a significant decrease in Vh:Cd and/or significant increase in IELs, we are left with data for ten patients. We will use this data in our simulations in Section 2.6.4.

Increased IELs, as seen in patients with CeD, correspond to 30 IELs per 100 enterocytes or more, while a healthy subject will have less than 30 IELs per 100 enterocytes. Lähdeaho et al. [36] measured the amount of IELs prior to and following the gluten challenge in terms of cells/mm. Our variable $I(t)$ corresponds to the amount of activated IELs at any given time. Prior to the challenge, there was a range of $23 - 146$ cells/mm and at the final time, $47 - 271$ cells/mm. For simplicity, we assume that this change in IEL density corresponds to the number of activated IELs. Since a healthy subject does not have any immune response, we set the number of activated IELs to zero for all time. Similarly, if a patient with CeD remains on a strict GFD, they will also have no immune response and thus no activated IELs. Thus, unless otherwise stated, the initial amount of activated IELs in simulations will be set equal to zero.

Recall, there are three parameters we need to estimate: d_1 and d_2, the destruction rates of the villus cells in the base and tip, respectively, by activated IELs, and p, the additional proliferation of crypt cells in the presence of activated IELs. We will calculate these values using the data from Lähdeaho et al. [36] in Section 2.6.4.

2.6.4 Calibrated Model Simulations in Disease Scenarios

In the previous simulations, we made assumptions to be able to simulate a theoretical patient with CeD, including the values for the three tuning constants in the activation function $A(t)$ and the parameters d_2 and p relating to the immune response. We would like to have better estimates of these values to have a more informed idea of their magnitude and value. In order to determine more accurate values, we simulate patients with CeD undergoing a gluten challenge using the data from the ten patients with significant decline in Vh:Cd and significant increase in IELs presented in Table 2.6 from the study by Lähdeaho et al. [36].

We first estimate the three tuning constants, a_1, a_2, and a_3, included in the activation function (2.25). At the baseline of the study, all patients had been strictly adhering to a long-term GFD and had a healthy Vh:Cd with "normal" IEL densities. Since everyone, healthy subjects and patients with CeD, has *non-activated* IELs and the model only considers the amount of *activated* IELs, we assume that the change from the initial to final IELs densities corresponds to the number of activated IELs at the final time with zero activated IELs at the initial time. While it could be that a patient with CeD might have some activated IELs, we assume that each patient was strictly adhering to the GFD and thus there would be negligible activated IELs.

We estimate the three parameters of the activation function $A(t)$, in order to achieve the number of activated IELs at the final time. In order to do so, appropriate ranges for each parameter were selected and multiple triples were used to estimate how close the simulated activated IELs at the final time were to the change in starting and ending values of IELs from data. Ranges and number of choices for each tuning constant are given in Table 2.7. In order to choose the appropriate set of tuning constants for varying patients, we conducted an extensive search. A total of 950 triples were generated to determine which triples produced the closest match to the data with respect to the final simulated number of IELs. Later, we will see that there is not one unique activation function to be used for each patient. Rather, we will consider how altering the parameters of the activation function effects the time courses of model variables.

The first parameter, a_1, determines the magnitude of the peak of activation. For now, we assume that this value falls between 0.1 and 1 and choose ten evenly-spaced values. Next, a_2 gives that time in minutes at which the peak activation will be reached. We assume that the IELs reach their peak activation between three and 21 hours. For simplicity, we only consider the peak activation at the top of the hour. Finally, a_3 determines the length of the period of the activation function; that is, the width of the non-negligible part of activation. A total of five values equally spaced between 100 and 14400 were chosen.

Beginning with Patient 1 from Lähdeaho et al. [36] included in Table 2.6, we have a patient that underwent a 29 day gluten challenge, that is, consumed gluten once a day for 29 days, with 55 activated IELs at the final time. We assume that this patient consumed

Table 2.7: Considered ranges for tuning constants for activation function $A(t)$.

Constant	**Range**	**Number of Choices**
a_1	$0.1 - 1$	10
a_2	$180 - 1260$	19
a_3	$100 - 14400$	5

gluten at time $t = 0$ and each of the 28 subsequent gluten intakes were spaced exactly 24 hours (1,440 minutes) apart. Thus, the beginning of day 30 corresponds to the end of day 29.

To determine how close the simulated activated IELs are to the data activated IELs, we use the objective function

$$F_1(a_1, a_2, a_3) = \frac{|I_f - I(t_{end})|}{|I_f|} \tag{2.26}$$

where I_f is the data for the number of activated IELs at the final time and $I(t_{end})$ is the simulated number of activated IELs at the final time. Smaller values of the objective function correspond to a better "fit" of the data. We use a threshold of 10% for relative error.

For Patient 1, of the 950 triples, 67 provided an objective function values less than 0.10. The smallest objective function value was approximately 0 which corresponded to 55.07 simulated activated IELs at the final time. The parameter values that provided this result are $a_1 = 0.40$, $a_2 = 720$, and $a_3 = 10,825$ with area under one simulated day of $A(t)$ as 73.76. Since the data corresponds to a daily gluten challenge, this activation function is pulsed every day, that is, every 1,440 minutes.

Clearly, there is not a unique set of triples to be used in $A(t)$ for each patient. A large part of this follows from the lack of data for each patient between the initial and final times. In order to consider varying scenarios, three tuning constant triples are chosen for each patient. This allows us to consider varying forms that still lead to the number of activated IELs at the final time. Table 2.8 provides three activation function choices that provide the appropriate number of IELs at the final time for the first three patients with significant changes in Vh:Cd and IELs as discussed in Section 2.6.3 from Lähdeaho et al. [36]. The three activation functions chosen for each of these patients is meant to cover a large space with varying a_1, a_2, and a_3 as well as area under $A(t)$ for the first 24 hours while focusing on the best "fit" as given by the objective function value.

Finally, we have three outstanding parameter values that need to be estimated from model output - d_1, d_2, and p. All three of these parameters are in regards to how activated IELs after the villus and crypt cell populations and not much is known regarding these three rates in the literature. As discussed earlier, we will initially assume that $d_1 = 0$ for a patient

Table 2.8: Potential activation function choices for activation function $A(t)$ for three patients from Lähdeaho et al. [36]. The area under $A(t)$ is integrated over the first day, corresponding to the level of activation per day.

Patient	a_1	a_2	a_3	**Area under** $A(t)$
1a	0.40	720	10,825	73.76
1b	0.30	960	10,825	55.32
1c	0.50	720	7,250	75.46
3a	0.80	720	3,675	85.96
3b	0.50	660	10,825	92.21
3c	1.0	180	7,250	150.71
4a	0.60	600	3,675	64.47
4b	0.50	300	10,825	92.20
4c	0.90	240	3,675	96.70

with CeD. That is, the activated IELs cause no change in the base of the villus. In order to estimate d_2 and p, we will again consider estimating these values on the patient level. Using the three activation functions per patient listed in Table 2.8, we can now use the Vh:Cd at the initial and final times for each patient to estimate these two remaining parameters.

We also would like to relate these cellular counts to the diagnostic factor Vh:Cd. Since

$$\text{Vh:Cd} = \frac{\alpha(V_1 + V_2)}{\beta C} \tag{2.27}$$

in our model structure, we would need to be able to calculate α and β for each patient. In order to simplify the number of calculations, we set $\delta := \alpha/\beta$. Each patient simulation will have its own δ, calculated as

$$\delta = \frac{C_0 \cdot \text{Vh:Cd}_0}{V_{1_0} + V_{2_0}} \tag{2.28}$$

where Vh:Cd_0 is the initial Vh:Cd for each patient from data and V_{1_0}, V_{2_0}, and C_0 are calculated based on parameter values as explained in Section 2.6.2.

We aim to ensure that the simulated average Vh:Cd over the last 24 hours approximates the final Vh:Cd from the Lähdeaho et al. [36] data. To determine how close these two values are, we use an objective function similar to the one we used earlier to estimate the activation function parameters

$$F_2(d_2, p) = \frac{|\text{Vh:Cd}_f - \text{Vh:Cd}(t_{avg-end})|}{|\text{Vh:Cd}_f|} \tag{2.29}$$

where Vh:Cd_f is the data for the final Vh:Cd and $\text{Vh:Cd}(t_{avg-end})$ is the average of the Vh:Cd over the last 24 hours of simulation. Smaller values of the objective function correspond to better fits of the data. We place a threshold for our relative error to be no more than 15%.

We begin by choosing ten of both d_2 and p in the ranges from 0.00001 and 0.0001, respectively, to 0.5 times the corresponding growth rates, r_2 and r_C, respectively, providing 100 potential sets. In the earlier theoretical simulations, we saw that values for d_2 and p approximately equal to r_2 and r_C, respectively, provided cellular growth that was not biologically feasible. Conversely, when d_2 and p were set to be 10% of r_2 and r_C, results made more sense. By beginning with this wide range, we hope to narrow down to potentially feasible values.

For the first activation function choice for Lähdeaho Patient 1, we find that of the 100 sets, there was one set of parameters that gave an objective function value of 0.15 or less. The parameter values are $d_2 = 0.000067$ and $p = 0.00010$. For the second activation function choice for Patient 1, we find two sets of parameters that satisfy our objective function threshold. For the third activation function choice for Patient 1, we find one set of parameters that satisfy our objective function threshold. We continue this same process with Patients 3 and 4. Sets of parameters are listed in Tables 2.9 and 2.10.

Multiple lines indicate multiple sets of potential parameters for the given patient and activation function choice. The only patient and activation function set which had no set of parameters that within our threshold was Patient 3, activation function b.

Considering these values, we find that d_2 can range between 0.00001 and 0.00052 while p ranges between 0.00010 and 0.0071. Since these values are not known in literature, this provides a feasible range to aid simulations. Note that the values we chose for the low immune response during the simulations for a theoretical patient with CeD in Section 2.6.2 are within these ranges.

2.7 Future Work

In order to accurately model cellular turnover in the small intestine, a narrow viewpoint of the immune response was needed since there are many unknowns in the dynamics and parameter values of this response. In future work, we aim to include a more detailed immune response. Demin et al. [21] developed a model that could describe a more detailed immune response necessary to test the efficacy of particular drug treatments, but their work could not be related to Q-MARSH. In conversations with gastroenterologists, they have stressed the advantages of being able to consider the dynamics of the Vh:Cd over time. In order to focus on the dynamics of histologic measurements, we needed to start with a simple immune response in order to correctly consider cellular turnover. Now that we have shown that the cellular turnover in our model accurately depicts the dynamics in healthy subjects and patients with CeD, we can begin to include more detail in the immune response.

Table 2.9: Fitted parameter values for d_2 and p for Patients 1 and 3 from Lähdeado et al. [36] ($d_1 = 0$). Multiple lines indicate multiple sets of potential parameters for the given patient and activation function choices.

Patient	d_2	p
1a	0.000067	0.00010
1b	0.000010	0.0015
	0.000067	0.00010
1c	0.000067	0.00010
3a	0.000010	0.00010
3b	0.000010	0.00010
3c	—	—

Table 2.10: Fitted parameter values for d_2 and p for Patient 4 from Lähdeado et al. [36] ($d_1 = 0$). Multiple lines indicate multiple sets of potential parameters for the given patient and activation function choices.

Patient	d_2	p
4a	0.00001	0.0057
	0.00001	0.0071
	0.000067	0.0029
	0.00012	0.0029
	0.00018	0.0029
	0.00024	0.0029
	0.00029	0.0029
	0.00035	0.0029
	0.00041	0.0029
	0.00047	0.0029
	0.00052	0.0029
4b	0.000010	0.0043
	0.000029	0.0015
	0.00035	0.0015
	0.00041	0.0015
	0.00047	0.0015
	0.00052	0.0015
4c	0.00001	0.0043
	0.00018	0.0015
	0.0002	0.0015
	0.00029	0.0015
	0.000035	0.0015
	0.00041	0.0015
	0.00047	0.0015
	0.00052	0.0015

In this chapter, we investigated the impact of different activation rate functions, all derived from the same form, but with varying tuning constants. In the future, we would like to investigate the impact of various IEL activation functions with different tuning constants and of different forms on both disease progression and recovery. Not much is know exactly how the immune response is activated, but our mathematical model has the ability to help fill in the details regarding the rate and form of activation.

Similarly, not enough information is known regarding the rates at which the immune response makes changes to the cells in the crypt and villus. Our work aimed to narrow potential ranges for the immune response parameter values from limited data. With more data points, it would be feasible to calculate these values more accurately and explicitly. We aim to continually improve this range of values calculated in Section 2.6.4 as well as validate these values using additional data.

Finally, by investigating appropriate sensitivity analysis methods, it would be possible to explore intervention strategies for promoting damage recovery. By determining which parameters are driving the system, those parameters could be varied to determine how altering their rates causes both short- and long-term changes to the system. Depending on the results of the simulations, therapeutics could be targeted to find a treatment for CeD.

2.8 Conclusions

In autoimmune diseases, the immune system mistakes part of the body as a foreign substance that needs to be killed and immunologists are unsure of why a person's genetics could dictate this destruction. Mathematics, including modeling techniques and analysis, has the potential to offer suggestions as to why these responses occur. Mathematicians have not readily modeled celiac disease, making this work novel and unique.

There are many key components in any immune response. In order to include every cytokine, antigen, and protein involved in the immune response of CeD, a relatively large system of differential equations would be required. However, the relation of many compartments can complicate a model and make it difficult to parameterize. Since most parameters involved in the immune response do not have known rates, it makes sense to

limit the number of parameters in the model by reducing the immune response to one compartment. With additional data sources, more detail could be included for a more accurate description of the immune response.

This work was based on limited data to estimate the parameter values related to the immune response. The Vh:Cd and IEL data from Lähdeaho et al. [36] were only collected at the initial and final time of the gluten challenge and the duration of the challenge varied among patients. While there are other shorter gluten challenge studies with associated data in the literature, we cannot couple those data points with the longer challenges as symptoms and histologic changes vary among patients. Ideally, instead of only the initial and final biopsies, data would be collected more frequently to help aid in parameterizing the model. It is likely that this data is not collected more frequently within gluten challenge studies due to the invasiveness of an upper endoscopy and biopsy in addition to patient unwillingness. It is our hope that as capsule endoscopies become more widely used that more data can also be collected which would aid our work.

It was our aim to estimate the immune-related parameters to better inform the reactions involved in the immune response of CeD. Knowing the rates of these interactions can better determine potential therapeutics to mitigate the negative effects of CeD. However, in order to be able to accurately predict efficient drug therapies, more detail will need to be included in the mathematical model.

Our simulation results showed that it takes approximately 18 days following activation of the immune response for the gut to return to "normal." This falls on the low end of the range estimated for patients with CeD following one gluten incident and indicates that our results may not be applicable to all patients suffering from CeD. Additionally, we found that it may be likely that the amount of gut damage with daily immune activation may follow a cyclic form starting as early as day three. If this is the case, gluten challenge studies to test potential drug treatments could be shortened significantly.

While the mathematical model considers the dynamics of gut health for a patient suffering from CeD, one should not forget the long-term effects of the disease, including a potential progression to cancer, the social consequences of needing to maintain a GFD, nor the anxiety that patients with CeD face every meal. A drug treatment to limit and stop the immune

response would help avoid not only the potential progression of the disease, but also remove the limitations that celiacs encounter daily. A mathematical model has the potential to help target potential therapeutics and finally find a drug treatment for CeD.

Chapter 3

Modeling the Effect of Environmental Transmission of *Clostridioides difficile* in Healthcare Settings

3.1 Introduction

Clostridioides difficile, formerly *Clostridium difficile* is an anaerobic, gram-positive bacillus spread among humans via the fecal-oral route [39]. *C. difficile* is the leading cause of infectious diarrhea and the most frequently identified healthcare-associated infection in United States hospitals [41]. *C. difficile* causes 223,900 cases in hospitalized patients and 12,800 deaths annually and the corresponding attributable healthcare costs is about one billion dollars [13]. In 2019, the Centers for Disease Control and Prevention (CDC) updated their analysis of the burden of the top antibiotic-resistant threats in the United States, classifying the threats according to their current and projected health and economic impacts. *C. difficile* was classified at the highest rank of 'Urgent Threat,' requiring immediate and aggressive public health action [13].

C. difficile produces endospores that are able to survive in harsh conditions for several weeks to months and are resistant to most commonly-used disinfectants and antimicrobial drugs [58]. When ingested, spores pass through the acidic environment of the stomach,

reaching the intestines. Spore germination is facilitated by primary bile acids produced by the liver and released in the intestines, but inhibited by secondary bile acids produced by gut bacteria that transform primary bile acids [58]. In individuals with antibiotic-induced changes in their gut microbiota, the decreased competition with other bacteria for resources and the lower levels of secondary bile acids promote spore germination and subsequent *C. difficile* growth. *C. difficile* produces two main toxins, TcdA and TcdB, that cause colonic inflammation and are responsible for clinical symptoms associated with infection that range from mild diarrhea to severe colitis and sepsis. However, if the host is able to mount an appropriate immune response against toxins TcdA and TcdB, they do not experience symptoms of infection. Since the prevention of *C. difficile* colonization in the colon is primarily due to the presence of a diverse gut microbiota, recent antibiotic use is the principle risk factor for colonization and infection [40].

Colonized patients, both symptomatic and asymptomatic, shed *C. difficile* spores that can survive for long periods on fomites, objects or surfaces that can harbor infectious agents [26, 34]. Evidence suggests that transmission pathways of *C. difficile* infection include interacting with fomites in healthcare settings [3, 23, 47, 48, 49, 56, 65, 66]. Transmission pathways also include the acquisition of the pathogen by occupying a contaminated room previously used by a *C. difficile*-infected patient and through the interaction with contaminated equipment or contaminated healthcare personnel who had previous contact with contaminated fomites and/or colonized patients.

Environmental control strategies include terminal cleaning and regular disinfection of isolation rooms or all rooms, and generally involve disinfectant substitutions, enhanced cleaning and disinfection practices, and recently, the use of automated disinfection technology. The CDC recommends daily cleaning of the immediate vicinity around an infected patient and disinfecting of an infected patient's entire room upon discharge with a chlorine-derived disinfectant, the most effective chemical against *C. difficile* spores [28, 43, 53]. In this study, it is important to note the difference between the terms 'clean' and 'disinfect.' Clean is defined as the removal of organic debris using vigorous scrubbing, while disinfect is the inactivation of pathogens. Compliance with these guidelines is often sub-optimal [9, 55] and

chlorine-derived cleaners are only used when necessary since they can damage and harm equipment.

Since *C. difficile* spores can survive for long periods on fomites not adequately cleaned or disinfected and are correlated with infection rates in healthcare settings, fomites can be a continuous source of transmission. This raises the question of what type of fomite in a healthcare setting contributes more to the transmission of *C. difficile* infection, fomites that are handled more frequently and tend to be cleaned and disinfected more often, or those that are handled less frequently and tend to be cleaned and disinfected less often. In most mathematical models of healthcare-associated infections, transmission is assumed to be direct, that is, through physical contact between susceptible and infectious individuals. Several studies have formulated agent-based [52], stochastic differential equation (SDE) [60], and ordinary differential equation (ODE) [67] models to incorporate environmental pathways in the transmission of healthcare-associated infections, but have not considered explicitly including the bacterial spore population. The agent-based models for *C. difficile* in hospital wards in Stephenson et al. [62] and Bintz et al. [8] included the level of pathogen implicitly as the contamination level of rooms. The stochastic queueing model in Chen et al. [16] represented in-hospital contact processes between environments and hosts and potential nosocomial pathogen transfer. Huang et al. [30] formulated ODE and SDE models to describe the transmission of methicillin-resistant *Staphylococcus aureus* (MRSA) among patients in a healthcare setting and explicitly tracked the density of bacteria in the environment, but did not consider different types of fomites as environmental reservoirs.

In this work, bacteria density is tracked on fomites that are classified as one of two types: high-touch, which are handled more often and therefore receive extra cleaning, and low-touch fomites, which are handled less often. Some examples of high-touch fomites include bed rails, tray tables, IV poles, doorknobs, sinks, and toilets, and some examples of low-touch fomites include curtains, furniture, thermostats, soap and paper towel dispensers, and trash cans [31]. These two environmental reservoirs of *C. difficile* are explicitly included in an ODE model to investigate the relative contribution of each type of fomite on new cases of *C. difficile* colonization among patients in a hospital ward. Since hospital wards have small patient populations, a stochastic simulation of the ODE model is used to quantify

the variations in possible model outcomes. Further stochastic simulations, driven by global sensitivity analysis results, determine the factors that increase *C. difficile* colonization and infection among patients and identify feasible strategies to prevent disease transmission. This work is in collaboration with Lindsey Fox (Eckerd College), Hannah Ritchie (North Carolina State University), Cristina Lanzas (North Carolina State University), Suzanne Lenhart, and Judy Day.

3.2 Model Formulation

The model in this study extends the system of ordinary differential equations formulated by Lanzas et al. [37], which determined that asymptomatic patients are a reservoir for *C. difficile*, by including environmental reservoirs. The dynamics of transmission are modeled using a system of six ODEs that describe the transitions between four patient classes and two environmental reservoir classes, summarized in Table 3.1. A schematic of the model is given in Figure 3.1 and the ODE model is given by the system in (3.1). However, since hospital wards have small population sizes whose dynamics may be sensitive to stochasticity, the average behavior of the patient population described by the deterministic simulation of the ODE model may mask extreme trajectories of the system that could occur. Therefore, this study stochastically simulates the ODE model using the Gillespie Stochastic Simulation Algorithm (GSSA) [27] to investigate the relative contributions of high-touch and low-touch fomites in a hospital ward on new cases of *C. difficile* colonization among patients. The GSSA events are explained in Section 3.2.3.

Because antibiotic use is the most significant risk factor for *C. difficile* colonization, resistant individuals (R) are defined as individuals that have not received recent antibiotic treatment and therefore have a normal, healthy gut microbiota and resistance to *C. difficile* colonization. Similarly, susceptible individuals (S) are defined as individuals that have recently received antibiotic treatment and therefore have a compromised gut microbiota and are susceptible to *C. difficile* colonization. Since a patient's microbiota returns to normal after they have ceased taking antibiotics, susceptible individuals can return to the resistant class. If a susceptible individual is exposed to *C. difficile* spores, they may become colonized.

Table 3.1: Patient and environmental reservoir classes for *C. difficile* model.

Variable	**Description** $[Units]$
R	Resistant $[individuals]$
S	Susceptible $[individuals]$
C	Colonized (asymptomatic) $[individuals]$
D	Diseased (symptomatic) $[individuals]$
P_H	*C. difficile* spore density on high-touch fomites $[spores \cdot cm^{-2}]$
P_L	*C. difficile* spore density on low-touch fomites $[spores \cdot cm^{-2}]$

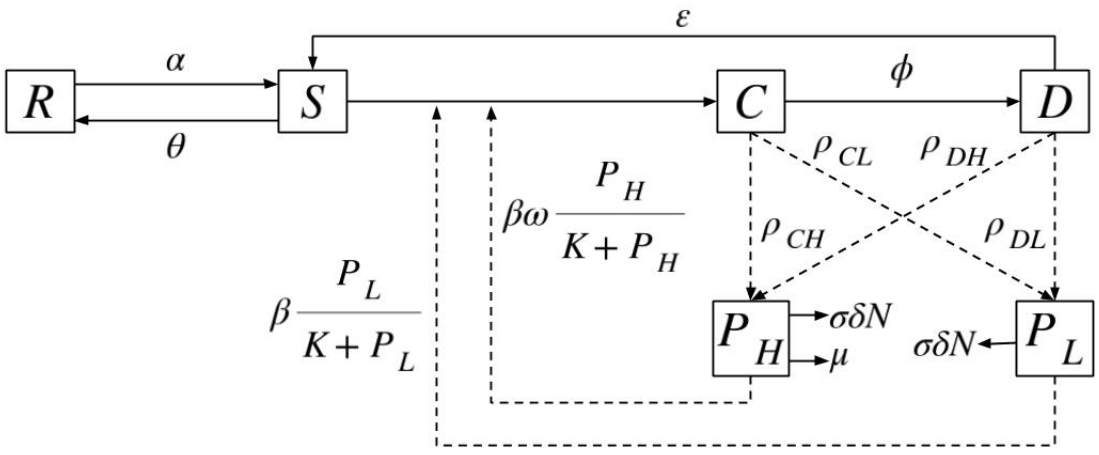

Figure 3.1: Schematic of *C. difficile* transmission. Admission and discharge rates for each patient class are not shown.

Colonized individuals (C) are defined as individuals that have *C. difficile* present in the gut, but are not exhibiting symptoms; while diseased individuals (D) are defined as colonized individuals who are symptomatic. Since diseased individuals receive antibiotic treatment for *C. difficile* infection, they return to the susceptible class if the treatment is successful; otherwise, they remain in the diseased class.

Lanzas et al. [37] differentiated between colonized individuals with and without an appropriate immune response against *C. difficile* infection before becoming symptomatic. This model does not make that distinction since the goal of this work is not to evaluate the contribution of asymptomatic patients to the transmission of *C. difficile*, as was the goal of Lanzas et al. [37]. Instead, this work investigates the relative contribution of two types of environmental reservoirs, to which asymptomatic and symptomatic patients contribute, in the transmission of *C. difficile*. Colonized patients both with and without an appropriate immune response against *C. difficile* infection shed spores that contribute to the bacterial spore population, and it is assumed that both shed spores at the same rate. Instead of differentiating between these colonized individuals, the transition rate from the colonized class to the diseased class incorporates the fraction of colonized individuals who do not mount a sufficient immune response against the toxins produced by the bacteria. The remaining colonized individuals are assumed to have mounted a sufficient immune response and stay in the colonized class for the duration of their hospital stay.

Unlike previous models of *C. difficile* spread, this model incorporates two classes for the environmental reservoirs of *C. difficile* in addition to the patient population: spore density on high-touch (P_H) and low-touch (P_L) fomites. Patients interact with high-touch fomites more often which are assumed to be cleaned regularly, as well as disinfected when a patient is discharged. Patients interact with low-touch fomites less often which are assumed to be disinfected only when a patient is discharged. (It is important to note the difference between the terms 'clean' and 'disinfect' described in Section 3.1.) Contamination of fomites occurs when spores are deposited on fomites through direct contact with soiled hands of colonized or diseased patients. Since *C. difficile* spores can survive on fomites for long periods of time, it is assumed in the model that natural spore death is negligible and spores are only removed through cleaning or disinfecting. If a susceptible individual is exposed to *C. difficile* via

contact with a fomite, the possibility of colonization depends on the type of fomite contacted and the spore density on that fomite. This could also depend on the duration of a contact with a fomite which is not considered in this model.

3.2.1 ODE Model

The dynamics of transmission illustrated in Figure 3.1 can be described using the following system of ODEs:

$$\begin{aligned}
R' &= a_R\delta(t)N - (k_R+\alpha)R + \theta S \\
S' &= a_S\delta(t)N + \alpha R - (k+\theta)S - \beta\left(\omega\frac{P_H}{K+P_H} + \frac{P_L}{K+P_L}\right)S + \varepsilon D \\
C' &= a_C\delta(t)N + \beta\left(\omega\frac{P_H}{K+P_H} + \frac{P_L}{K+P_L}\right)S - (k+\phi)C \\
D' &= a_D\delta(t)N + \phi C - (k_D+\varepsilon)D \\
P_H' &= \rho_{CH}C + \rho_{DH}D - (\sigma\delta(t)N + \mu)P_H \\
P_L' &= \rho_{CL}C + \rho_{DL}D - \sigma\delta(t)NP_L
\end{aligned} \tag{3.1}$$

Patients may be admitted into or discharged from any patient class. Discharge rates account for release from the hospital, transfer to another ward, and death. When patients are discharged, new patients are admitted to all four patient classes according to admission proportions a_R, a_S, a_C, and a_D to maintain a constant ward population, N. The rate at which patients are admitted into a patient class also depends on the total discharge rate, $\delta(t)$. The total ward population, N, is the sum of all patient classes and is assumed to be constant at hospital ward capacity. This assumption, with the fact that the sum of the admission proportions a_R, a_S, a_C, and a_D is 1, leads to the following expression for the total discharge rate,

$$\delta(t) = \frac{1}{N}\Big(k_R R(t) + k\big(S(t)+C(t)\big) + k_D D(t)\Big). \tag{3.2}$$

The discharge rates of the patient classes, k_R, k, and k_D, and the total ward population, N, are constant, while the populations of the patient classes depend on time. Thus, $\delta(t)$

also depends on time and has units of day^{-1}. Initial conditions of the patient classes are also determined by the admission proportions. The initial condition of patient class i is $a_i N$ rounded to the nearest non-zero integer, for $i = R, S, C, D$. Model parameters are described in Table 3.2; details for parameter values and their calculations are included in Section 3.2.2.

Note that the basic reproduction number, $\mathcal{R}_0$, cannot be calculated since a disease-free equilibrium does not exist. In this model, the basic reproduction number would define the average number of secondary colonizations produced by a singular *C. difficile* colonization, C or D patient, in an otherwise *C. difficile*-free hospital ward. At a disease-free equilibrium, there would be an absence of colonization and pathogen, that is, $C = D = P_H = P_L = 0$. The system in (3.1) with this absence does not give a non-trivial, biologically feasible, disease-free equilibrium point satisfying the constraint that the total ward population, N, remains constant.

3.2.2 Model Parametrization

Patient Class Parameters

Many parameter values in this study are the same as or calculated from parameter values in Lanzas et al. [37], who determined their parameter values from data collected retrospectively from Barnes-Jewish Hospital in St. Louis, Missouri in 2008.

The admission proportions, a_R, a_S, a_C, and a_D, are taken from [61], who updated the values from [37] based on later research supporting a higher admission proportion of asymptomatically colonized individuals. The discharge rates, k_R, k, and k_D, are taken from [37] and are the inverse of patients' average duration in the hospital. The discharge rates for the susceptible and colonized classes are assumed to be the same, $k_S = k_C = k$, since asymptomatic patients are typically not tested for *C. difficile* and therefore cannot be differentiated from susceptible patients. Additionally, colonized patients do not require treatment for *C. difficile* infection and do not need to remain in the hospital longer than susceptible patients due to *C. difficile* infection. The antibiotic prescription rate, α, is also taken from [37] and is based on admission rates and the percentage of individuals who received antibiotics during their stay.

Table 3.2: Description of *C. difficile* model parameters.

Parameter	Description [$Units$]	Reference Value	Source
a_R	proportion of individuals admitted into R [$dimensionless$]	0.75	[61]
a_S	proportion of individuals admitted into S [$dimensionless$]	0.09	[61]
a_C	proportion of individuals admitted into C [$dimensionless$]	0.15	[61]
a_D	proportion of individuals admitted into D [$dimensionless$]	0.01	[61]
k_R	discharge rate of R [day^{-1}]	0.33	[37]
k	discharge rate of S and C [day^{-1}]	0.15	[37]
k_D	discharge rate of D [day^{-1}]	0.068	[37]
α	antibiotic prescription rate [day^{-1}]	0.5	[37]
θ	gut restoration of colonization resistance rate [day^{-1}]	0.033	[51]
ε	successful treatment of infection rate [day^{-1}]	0.08	[46]
ϕ	disease rate of colonized individuals with insufficient immune response [day^{-1}]	0.024	[15, 37]
ρ_{CH}	shedding rate of C into P_H [$spores \cdot cm^{-2} \cdot individuals^{-1} \cdot day^{-1}$]	0.057	[17, 33, 54]
ρ_{CL}	shedding rate of C into P_L [$spores \cdot cm^{-2} \cdot individuals^{-1} \cdot day^{-1}$]	0.029	[17, 33, 54]
ρ_{DH}	shedding rate of D into P_H [$spores \cdot cm^{-2} \cdot individuals^{-1} \cdot day^{-1}$]	0.123	[17, 33, 54]
ρ_{DL}	shedding rate of D into P_L [$spores \cdot cm^{-2} \cdot individuals^{-1} \cdot day^{-1}$]	0.063	[17, 33, 54]
σ	proportion of P_H and P_L spores killed due to disinfection upon discharge [$individuals^{-1}$]	0.83	[55]
μ	rate of P_H spores killed due to extra cleaning [day^{-1}]	0.66	[55]
K	half-saturation constant [$spores \cdot cm^{-2}$]	7.5	[38]
β	colonization rate upon transfer of spores from a fomite [day^{-1}]	0.338	—
ω	weighting constant for high-touch fomites [$dimensionless$]	1.96	[17]

After ceasing antibiotic treatment, the gut returns to normal after approximately 30 days [51]. Thus, the rate at which the gut restores its resistance to *C. difficile* colonization, θ, is the inverse of the time to restoration. For approximately 80% of patients with *C. difficile* infection, symptoms resolve within a ten day treatment of antibiotics vancomycin or metronidazole [46]. Thus, the successful treatment rate, ε, is the product of the proportion of successfully treated patients and the inverse of the treatment duration. Approximately 60% of patients have an appropriate immune response to the toxins produced by *C. difficile* [35]. According to [15], those that do not mount an appropriate immune response start to display symptoms after an average of six days. Thus, the disease rate of colonized patients, ϕ, is the product of the complement of the proportion of patients with an appropriate immune response and the inverse of the incubation period.

Shedding Parameters

The shedding rates, ρ_{CH}, ρ_{DH}, ρ_{CL}, and ρ_{DL}, and the weighting constant for high-touch fomites, ω, are associated with the two environmental reservoir classes P_H and P_L. The four shedding rates, ρ_{ij}, where $i = C, D$ and $j = H, L$, are determined by multiplying the contacts per day per individual with a fomite, cpd_j, and the spores per square centimeter per contact, spc_i, as summarized in Table 3.3. The weighting constant for high-touch fomites, ω, is the ratio of the contacts per day per individual with high-touch fomites, cpd_H, to the contacts per day per individual with low-touch fomites, cpd_L.

An analysis by [54] cultured stool, skin, and environmental samples from patients with *C. difficile* infection before, during, and after treatment. They found an average of 3.333 spores on the hand of the individual making contact an asymptomatic patient and an average of seven spores after contact with a symptomatic patient. With the average surface area of a hand being 0.054 square meters [33], this results in 0.006 spores per square centimeter per contact with a colonized patient (spc_C) and 0.013 spores per square centimeter per contact with a diseased patient (spc_D).

A study by [17] observed the number of hand-touch contacts by patients with fomites. Over a period of 66 hours, they observed 12 patients make 470 contacts, 311 of which were with high-touch fomites and 159 were with low-touch fomites. Based on this data, patients

Table 3.3: Calculation of shedding parameters ρ_{ij} where $i = C, D$ and $j = H, L$.

Patient Class	**Environmental Reservoir**	**Spores per cm² per contact** ($\mathbf{spc_i}$)	**Contacts per day per individual** ($\mathbf{cpd_j}$)	**Shedding rate** $\rho_{\mathbf{ij}} = \mathbf{spc_i} \cdot \mathbf{cpd_j}$
C	P_H	$spc_C = 0.006$	$cpd_H = 9.424$	$\rho_{CH} = 0.057$
C	P_L	$spc_C = 0.006$	$cpd_L = 4.818$	$\rho_{CL} = 0.029$
D	P_H	$spc_D = 0.013$	$cpd_H = 9.424$	$\rho_{DH} = 0.123$
D	P_L	$spc_D = 0.013$	$cpd_L = 4.818$	$\rho_{DL} = 0.063$

have 9.424 contacts per day per individual with high-touch fomites (cpd_H), 4.818 contacts per day per individual with low-touch fomites (cpd_L), and are $\omega = cpd_H/cpd_L = 1.96$ times more likely to contact a high-touch fomite than a low-touch fomite.

Cleaning and Disinfecting Parameters

Based on CDC recommendations, it is assumed that both types of fomites are disinfected after a patient is discharged and only high-touch fomites receive extra, regular cleaning. However, since this is a continuous time model, these are not discrete disinfecting or cleaning events but per day rates.

Shams et al. [55] swabbed fomites to determine the typical microbial burden on high-touch fomites after extra cleaning and disinfecting upon discharge. They observed that after extra cleaning, 34% of sites tested contained pathogen and after disinfecting upon discharge, 17% of sites tested contained pathogen. Thus, for this model, the rate at which *C. difficile* spores are killed on high-touch fomites due to extra cleaning, μ, is 0.66 per day, meaning 34% of spores remain on fomites and 66% of spores are killed per day due to extra cleaning.

This model does not describe the behavior of individual patients and spores in individual rooms, rather, it describes average mass-action dynamics for homogeneously mixed classes of patients and fomites that contain spores. Thus, disinfection upon discharge, which would only kill spores on fomites in the discharged patient's room, needs to describe the proportion of all spores that are associated with a single patient and killed when disinfected. Therefore, the disinfection upon discharge rate is the product of the total discharge rate, $\delta(t)$, the total ward population, N, and the proportion of spores killed per individual discharged, σ. From [55], the proportion of spores killed per individual discharge is set to 0.83 per individual, meaning 17% of spores remain on fomites and 83% of spores are killed after disinfecting upon the discharge of an individual.

Force of Infection Parameters

For this model, the force of infection is the rate at which susceptible individuals become colonized with *C. difficile* after contacting a fomite and is given by the expression

$$\beta\left(\omega\frac{P_H}{K+P_H}+\frac{P_L}{K+P_L}\right). \tag{3.3}$$

It is the sum of the rates of colonization due to contact with a high-touch and low-touch fomite. The rate of colonization for each type of fomite is the product of the proportion of spores transferred from a fomite to a patient after a contact, $P_j/(K+P_j)$ for $j=H,L$, and the rate at which a susceptible individual becomes colonized by the transferred spores, β. The proportion of spores transferred for each type of fomite is assumed to have a type II functional response [29] with half-saturation constant K, where the proportion of spores transferred increases as the pathogen level on the fomite increases, but saturates at high levels of pathogen.

Since a high-touch fomite is defined as a fomite that is contacted more often, the colonization rate for high-touch fomites has an extra weighting parameter ($\omega > 1$) to account for the higher frequency of contacts. In [17], patients made 470 contacts with fomites in a hospital room, 311 of which were with high-touch fomites and 159 were with low-touch fomites. Therefore, a patient is $\omega = 1.96$ times more likely to contact a high-touch fomite than a low-touch fomite. It is important to note that this is the same study used to calculate the factors cpd_H and cpd_L which are used in the calculation of the shedding parameters.

In a study by [38], susceptible mice were held in cages contaminated with *C. difficile* spores isolated from a multi-hospital outbreak. They found that an average of 7.5 spores per square centimeter in a cage resulted in 50% of the mice in the cage becoming colonized. To estimate the value of β, it is assumed that the force of infection given by (3.3) is equal to 0.5 per day when pathogen levels on both types of fomites are 7.5 spores per square centimeter. This results in the linear relationship between K and β where

$$\beta = \frac{0.5\cdot(7.5+K)}{(\omega+1)\cdot 7.5}. \tag{3.4}$$

Assuming that when contacts transferring 50% of spores from a fomite to a patient results in 50% of patients becoming colonized per day, i.e., that K is 7.5 spores per square centimeter, then β will be 0.338 per day.

3.2.3 Stochastic Simulation of ODE Model Events

The deterministic ODE model given by the system in (3.1) generates the average behavior of the system. However, hospital wards generally have small population sizes whose dynamics can be sensitive to stochasticity; a particular outbreak of *C. difficile* in a hospital ward may not necessarily follow the behavior of an average outbreak. Thus, the events modeled by the ODE were simulated using the Gillespie Stochastic Simulation Algorithm (GSSA) [4, 6, 27] to understand the variability among different trajectories, or time courses, of the system.

The GSSA views interactions between populations as discrete events that occur randomly and uses a Markov Chain Monte Carlo method to sample from the solution at each time point, creating a possible time course of a system. Based only on the current population levels, not past levels, the time step and the event that occurs after that time step are determined probabilistically by propensity functions. Population levels are then updated by integer values based on the event that occurred [2, 6, 27].

For N population classes and M reactions, the state of the system at time t, given by $\mathbf{X}(t) = [X_1(t), X_2(t), \ldots, X_N(t)]$, is updated by the GSSA. The occurrence of reactions are assumed to follow a Poisson point process, meaning reactions occur at known, constant rates that are independent of the time since the last reaction. The probability density function of the Poisson distribution

$$P(x|\lambda) = \frac{\lambda^x}{x!}e^{-\lambda}, \qquad x \in [0, \infty) \tag{3.5}$$

describes the probability P of a reaction occurring within a given interval exactly x times given that the reaction occurs within the given interval λ times on average. The exponential distribution

$$P(\tau|\lambda) = \lambda e^{-\lambda\tau}, \qquad \tau \in [0, \infty) \tag{3.6}$$

describes the probability P of the time interval between reactions in a Poisson point process where the time interval to next reaction τ given the reaction occurs in the interval λ on average. The probability that any of the M reactions will occur in the time interval $(t, t + \tau)$, given the system is in the present state $\mathbf{X}(t)$, is the sum of the propensity functions $a_0 = \sum_{j=1}^{M} a_j$. The time to next reaction τ is chosen from the exponential distribution $P(\tau|a_0)$ [2, 19].

A random number r_1 is generated from a uniform distribution on $[0, 1]$ so that a value for τ can be calculated, given by

$$\tau = \frac{1}{a_0} \ln\left(\frac{a_0}{r_1}\right). \tag{3.7}$$

The probability that reaction $j = \{1, \ldots, M\}$ is the next reaction is a_j/a_0. If a random number r_2 generated from a uniform distribution on $[0, 1]$ is greater than $\sum_{k=1}^{j-1} a_k/a_0$ and less than or equal to $\sum_{k=1}^{j} a_k/a_0$, that will be the reaction that occurs next. This process continues for the pre-determined duration of the trajectory [2, 6, 27].

In this study, stochastic simulation via the GSSA shows particular time courses of the transmission of *C. difficile* by incorporating the probabilistic occurrence of interactions among the six state variables described in Table 3.1. The state variables are involved in 15 different events for this model as developed by Fox [25] and summarized in Table 3.4. The propensity function of an event is the product of the rate at which that event occurs on average and the size of the affected state variable at that time. For example, in the event that a resistant individual receives a prescription for an antibiotic, the propensity function is the product of the rate at which antibiotics are prescribed, α, and the size of the affected state variable, $R(t)$. When this event occurs, one resistant individual becomes susceptible and all other state variables are not changed. Note that the numbering of the events does not indicate that the events occur in that order; the ordering of the events occurs randomly as described earlier.

Since it is assumed that the total hospital ward population is held constant at capacity, admission into the ward can only occur if a patient is discharged, given by event **1**, **2**, **3**, or **4** in Table 3.4. The class to which a new patient is admitted is randomly chosen, weighted by the admission proportions. Thus, the state space may be updated in one of four possible ways. When a patient is discharged, both high-touch and low-touch fomites are disinfected. Thus, as a result of a discharge event, the spore density on high-touch and low-touch fomites is reduced by the proportion of spores killed when fomites are disinfected due to an individual discharge, σ. In the event that high-touch fomites receive extra cleaning, the spore density on high-touch fomites is reduced by the proportion of spores killed due to extra cleaning per

Table 3.4: GSSA events with state space $\mathbf{X}(t) = [R(t), S(t), C(t), D(t), P_H(t), P_L(t)]$. Key for events **1 - 4**: State space update if new patient is admitted into ${}^{*}R$, ${}^{**}S$, ${}^{\dagger}C$, or ${}^{\dagger\dagger}D$.

No.	Event	State Update	Propensity Function
1	discharged from R	$[0,0,0,0,-\sigma P_H,-\sigma P_L]^{*}$, $[-1,1,0,0,-\sigma P_H,-\sigma P_L]^{**}$, $[-1,0,1,0,-\sigma P_H,-\sigma P_L]^{\dagger}$, or $[-1,0,0,1,-\sigma P_H,-\sigma P_L]^{\dagger\dagger}$	$k_R R$
2	discharged from S	$[1,-1,0,0,-\sigma P_H,-\sigma P_L]^{*}$, $[0,0,0,0,-\sigma P_H,-\sigma P_L]^{**}$, $[0,-1,1,0,-\sigma P_H,-\sigma P_L]^{\dagger}$, or $[0,-1,0,1,-\sigma P_H,-\sigma P_L]^{\dagger\dagger}$	kS
3	discharged from C	$[1,0,-1,0,-\sigma P_H,-\sigma P_L]^{*}$, $[0,1,-1,0,-\sigma P_H,-\sigma P_L]^{**}$, $[0,0,0,0,-\sigma P_H,-\sigma P_L]^{\dagger}$, or $[0,0,-1,1,-\sigma P_H,-\sigma P_L]^{\dagger\dagger}$	kC
4	discharged from D	$[1,0,0,-1,-\sigma P_H,-\sigma P_L]^{*}$, $[0,1,0,-1,-\sigma P_H,-\sigma P_L]^{**}$, $[0,0,1,-1,-\sigma P_H,-\sigma P_L]^{\dagger}$, or $[0,0,0,0,-\sigma P_H,-\sigma P_L]^{\dagger\dagger}$	$k_D D$
5	antibiotic prescribed	$[-1,1,0,0,0,0]$	αR
6	gut restored against colonization	$[1,-1,0,0,0,0]$	θS
7	new C due to P_H	$[0,-1,1,0,0,0]$	$\beta\omega\frac{P_H}{K+P_H}S$
8	new C due to P_L	$[0,-1,1,0,0,0]$	$\beta\frac{P_H}{K+P_H}S$
9	symptoms developed	$[0,0,-1,1,0,0]$	ϕC
10	infection successfully treated	$[0,1,0,-1,0,0]$	εD
11	high-touch fomites cleaned	$[0,0,0,0,-\frac{\mu}{N}P_H,0]$	N
12	C sheds into P_H	$[0,0,0,0,spc_C,0]$	$cpd_H C$
13	C sheds into P_L	$[0,0,0,0,0,spc_C]$	$cpd_L C$
14	D sheds into P_H	$[0,0,0,0,spc_D,0]$	$cpd_H D$
15	D sheds into P_L	$[0,0,0,0,0,spc_D]$	$cpd_L D$

individual, μ/N. The extra cleaning event occurs at a rate of once per day, on average, per CDC guidelines and affects all patient state variables. Thus, the propensity function is N.

Shedding rates, ρ_{ij} where $i = C, D$ and $j = H, L$, are determined by multiplying the contacts per day per individual, cpd_j, by the spores per square centimeter transferred per contact, spc_i, as detailed in Section 3.2.2. Thus, the propensity function for a shedding event is the contacts per day per individual, $cpd_j, j = H, L$, multiplied by the number of individuals in the patient class of the shedding patient, either C or D. The pathogen density of the fomite shed on increases by the spores per square centimeter transferred per contact, $spc_i, i = C, D$.

3.3 Sensitivity Analysis

A global sensitivity analysis as described in Marino et al. [44] conducted on the ODE model determined the parameters that have the most significant impact on three outcomes of interest

1. the incidence of colonized patients due to contact with high-touch fomites,
2. the incidence of colonized patients due to contact with low-touch fomites, and
3. the incidence of diseased patients.

Parameter values were sampled using Latin Hypercube Sampling (LHS), Partial Rank Correlation Coefficients (PRCC) were computed as the sensitivity indices, and associated p-values were calculated to determine the significance of each PRCC value. The PRCC values and p-values for each parameter relative to the three outputs of interest aid in quantifying each parameter's contribution to each outcome. Monotonicity between parameters and the three outcomes was confirmed.

Several scenarios considering changes in the values of these impactful parameters were stochastically simulated to determine the factors that increase incidence. The results described in Section 3.4 show that increased contacts with high-touch fomites cause the largest increase in the incidence of colonizations and an increased rate at which colonized patients become diseased is responsible for the largest increase in the incidence of diseased. It

is important to note that from the formulation of the model, only the incidence of colonized patients can be linked with the environmental reservoir class that lead to colonization.

All parameters listed in Table 3.2 were included in the sensitivity analysis, except for the four admission proportions since they must sum to 1. To construct intervals for the LHS sampling, values for each parameter were randomly selected from a uniform probability distribution in the interval 50% below to 50% above the reference value. The parameter ω was an exception to this convention since ω must be at least 1 to account for the increased frequency of contacts with high-touch fomites. The parameter ω was sampled from the uniform probability distribution in the interval $(1, 2.92)$.

Since the admission proportions were not included in the sensitivity analysis, but may have an effect on the results, they were altered manually and results were compared to the case using baseline parameter values. The admission proportions were estimated from hospital data in [37]. While resistant and diseased individuals are easily identified, susceptible and colonized individuals are not easily differentiated without testing. As such, a_S and a_C are the only two admission proportions varied. For each outcome, two additional cases were considered where a_S was increased by 0.05 with a_C decreased by 0.05, and vice versa. Initial conditions are altered to match the new admission proportions.

For each output, 2,000 distinct parameter sets were chosen. PRCC values with a p-value less than 0.05 are considered to be statistically significant. Significant PRCCs for each parameter in each of the three outcomes are given in Table 3.5. A positive PRCC value indicates that an increase in the parameter value produces an increase in the outcome. A negative PRCC value indicates that an increase in the parameter value produces a decrease in the outcome [44]. Based on their p-values, all parameters have a PRCC value deemed significant in at least one of the three outcomes.

When the admission proportions a_S and a_C are varied, these results largely remain unchanged. However, when a_S is increased by 0.05 with a_C decreased by 0.05, the significant PRCC values for k_D and ε become negative for the incidence of diseased and are no longer significant for the incidence of colonized due to both high-touch and low-touch fomites. When a_C is increased by 0.05 with a_S decreased by 0.05, ρ_{DH} and ω no longer have significant PRCC values for the incidence of colonized due to low-touch fomites.

Table 3.5: Significant PRCC values for three outcomes for *C. difficile* model. A cell with no value indicates that the PRCC value did not have an associated p-value less than 0.05 and thus has no significant effect on an outcome.

Parameter	**Incidence of C due to high-touch fomites**	**Incidence of C due to low-touch fomites**	**Incidence of D**
k_R	-0.14	-0.23	0.61
k	-0.86	-0.88	-0.82
k_D	-0.18	-0.24	—
α	0.49	0.52	-0.11
θ	-0.14	-0.11	—
ε	-0.26	-0.21	—
ϕ	0.15	0.15	0.96
ρ_{CH}	0.77	0.07	0.33
ρ_{CL}	—	0.76	0.09
ρ_{DH}	0.49	0.07	0.17
ρ_{DL}	—	0.49	0.07
σ	-0.85	-0.88	-0.52
μ	-0.24	—	-0.13
K	-0.88	-0.88	-0.58
β	0.89	0.89	0.55
ω	0.87	0.15	0.43

Since all parameters were deemed significant in at least one of the three outcomes, further work was needed to obtain a ranking of the parameters for all three outcomes. Multiple pairwise comparison z-tests as outlined in Marino et al. [44] were used to determine if one parameter has a greater impact on an outcome than another, that is, whether there were statistical differences between corresponding PRCC values. The null and alternative hypotheses for each test is that the absolute value of two PRCC values are and are not equal, respectively. The significance level for each test is 0.05 divided by the number of parameters in the given set of z-tests. Failing to reject the null hypothesis means that it cannot be determined which of two parameters has a greater impact on an outcome. Rankings of the parameters with significant PRCC values for the three outcomes are given in Table 3.6. These rankings remain the same when a_S and a_C are varied.

The specific scenarios investigating altered parameter values considered for further stochastic simulations were largely driven by the results in Table 3.6; the highest ranking parameters in each of the three outcomes were increased by 50% of their baseline value in Table 3.2. Since the primary focus of this study is to understand how the two environmental reservoirs contribute to the transmission of *C. difficile* in a hospital ward, in some scenarios, multiple related parameters were increased. For example, the rate at which spores are shed onto fomites, ρ_{ij} where $i = C, D, j = H, L$, depends on how often the fomites are contacted, as described in Section 3.2.2. Thus scenarios increasing various shedding rates together, as well as with ω, were investigated. Other choices of scenarios, such as increased efficacy of cleaning and disinfecting, were not driven by sensitivity analysis results but instead by curiosity about their impact on the environmental reservoirs.

3.4 Simulation Results

3.4.1 Results of the Baseline Scenario

Using the set of baseline parameter values given in Table 3.2, the deterministic ODE model was simulated using MATLAB's ode45 solver [45]. Using the same parameter values, MATLAB was also used to apply the GSSA for the stochastic simulation of the model

Table 3.6: Ranking of significant PRCC values for parameters for three outcomes from most to least sensitive based upon pairwise z-test results. Multiple parameters listed in the same cell indicate that their PRCC values are not statistically different. Shaded cells indicate positive PRCC values; non-shaded cells indicate negative PRCC values.

	Incidence of C due to high-touch fomites	**Incidence of C due to low-touch fomites**	**Incidence of D**
most	β, ω	β	ϕ
	ρ_{CH}	ρ_{CL}	k_R
	α, ρ_{DH}	α, ρ_{DL}	β
	ϕ	ϕ, ω	ω
↓	k_R, k_D, θ	ρ_{CH}, ρ_{DH}	ρ_{CH}
↓	ε, μ	θ	ρ_{CL}, ρ_{DH}, ρ_{DL}
↓	k, σ	k_R, k_D, ε	α, μ
	K	k, σ, K	σ
			K
least			k

events. The deterministic and stochastic simulations show that on average approximately 77% of new colonizations are caused by contact with a high-touch fomite and 23% are caused by contact with a low-touch fomite, despite the extra cleaning high-touch fomites receive.

Incidence results of the deterministic and stochastic simulations are compared in Table 3.7 and plots comparing both are shown in Figure 3.2. Individual trajectories from the stochastic simulation are shown in Figure 3.3. As seen in Figure 3.2, the average of the trajectories for the patient and environmental reservoir classes seems to quickly settle on an endemic equilibrium. However, individual trajectories can largely vary from the average. In the 100 forty-day time courses of the stochastic simulation of the system, the incidence of colonized ranges from 14.74 to 38.72 per 1,000 admissions and the incidence of diseased ranges from 11.71 to 30.99 per 1,000 admissions within one standard deviation of the mean.

The deterministic simulation and average of the stochastic simulation trajectories match in incidence results and time course behavior, as seen in Figure 3.2. For this reason, all of the following results will be presented from the stochastic simulation only, to show the variability among *C. difficile* outbreaks in small simulated hospital wards. Also, in all scenarios described below, the stochastic simulation results consist of 100 trajectories of the system described by (3.1) and GSSA events described in Table 3.4 with parameter values from Table 3.2 unless otherwise stated and used a total hospital ward population of $N = 30$ patients over 40 days. The initial conditions for the patient classes were $R_0 = 21$, $S_0 = 3$, $C_0 = 5$, and $D_0 = 1$ individuals, as explained in Section 3.2.1. The initial conditions for the environmental reservoir classes are $P_{H_0} = P_{L_0} = 0.01$ spores per cm^2, an amount considered as a well-maintained environment, disjoint from the frequency that fomites are touched.

3.4.2 Results from Increasing Spores Shed on and Contacts with Fomites

Changes in the shedding rates, ρ_{CH}, ρ_{DH}, ρ_{CL}, and ρ_{DL}, and subsequent changes in the weighting constant of high-touch fomites, ω, produce the largest increase in the incidence of colonized patients and the contribution of high-touch fomites to new colonizations.

Table 3.7: Incidence results of the deterministic and stochastic simulations of the system given in (3.1) with baseline parameter values from Table 3.2. Results of the deterministic simulation are reported as new cases per 1,000 admissions. Results of the stochastic simulation are reported as the mean ± standard deviation of new cases per 1,000 admissions in 100 trajectories of the system.

	Deterministic Simulation	**Stochastic Simulation (Scenario #0)**
Incidence of C	28.09	26.73 ± 11.99
due to high-touch fomites	21.61	20.63 ± 10.20
due to low-touch fomites	6.48	6.09 ± 4.98
Incidence of D	21.69	21.35 ± 9.64

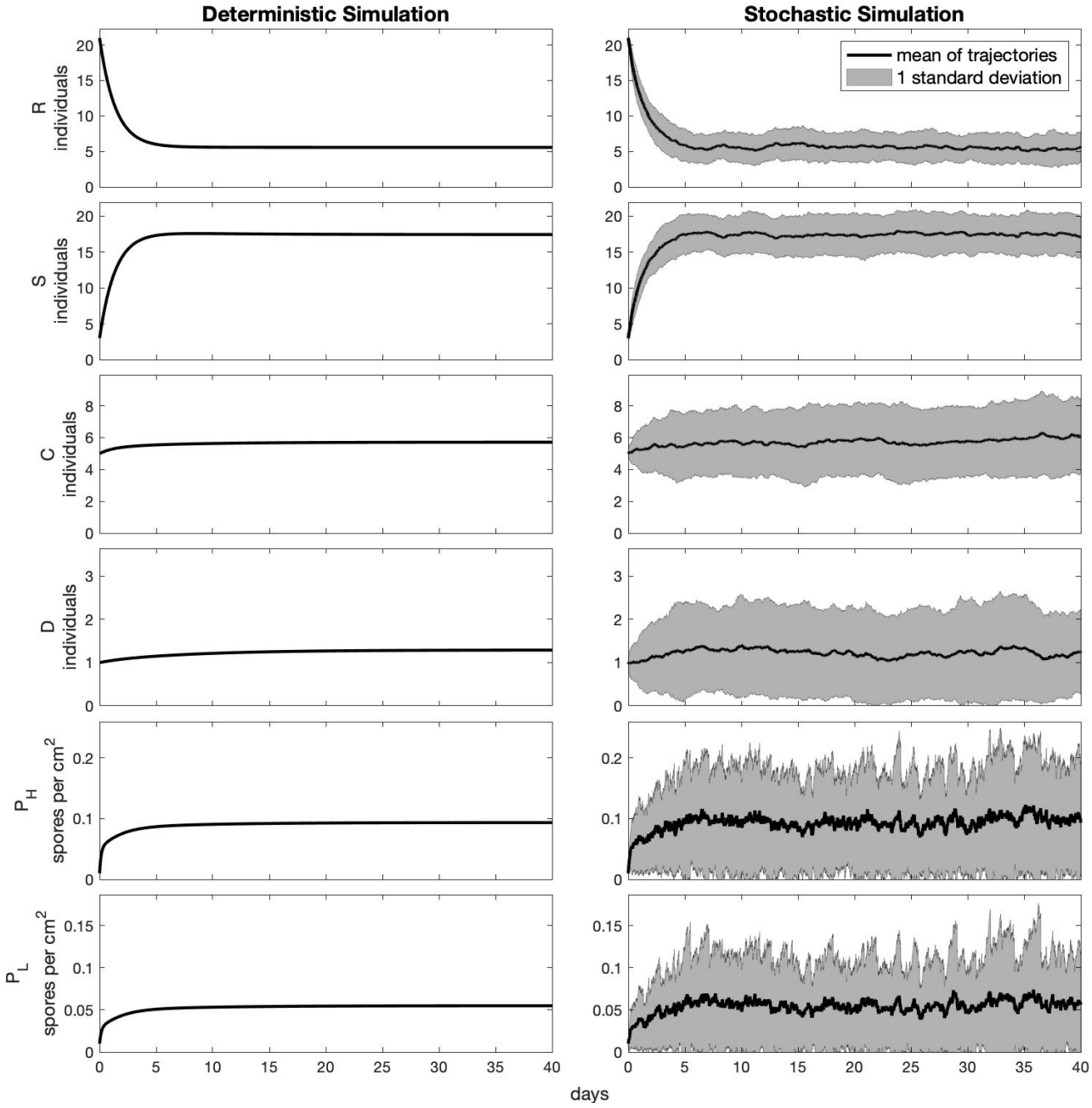

Figure 3.2: Plots of the deterministic (left column) and stochastic simulations (right column) of the system given in (3.1) over 40 days with baseline parameter values from Table 3.2 (Scenario #0). The thick black line in the plots in the right column shows the average value of the 100 trajectories at each time and the shaded area shows one standard deviation on either side of the average value.

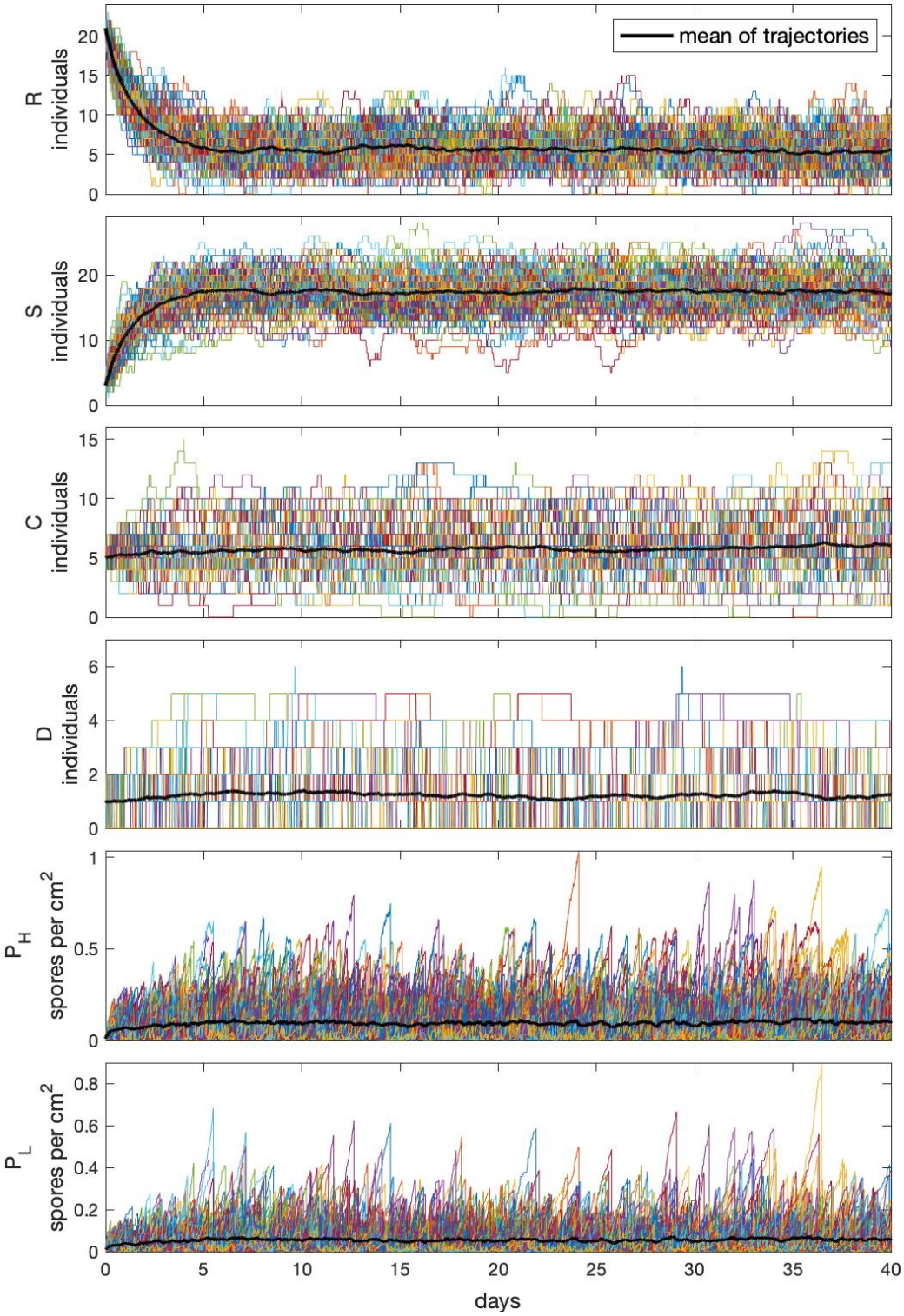

Figure 3.3: Plots of the 100 trajectories of the stochastic simulation of the system given in (3.1) over 40 days with baseline parameter values from Table 3.2 (Scenario #0). The thick black line shows the average value of the trajectories at each time.

The shedding rates onto high-touch fomites, ρ_{CH} and ρ_{DH}, are the second and third, respectively, most impactful parameters to the incidence of colonized due to high-touch fomites, as shown in Table 3.6. The shedding rates onto low-touch fomites, ρ_{CL} and ρ_{DL}, are the second and third, respectively, most impactful parameters to the incidence of colonized due to low-touch fomites, as shown in Table 3.6. Since the PRCC values for these parameters are positive, the sensitivity analysis indicates that an increase in any of these parameters is correlated with an increase in the respective outcomes. Since the shedding rates are determined by the product of the spores per square centimeter per contact, $spc_i, i = C, D$, and the contacts per day per individual, $cpd_j, j = H, L$, a 50% increase in one of these factors would produce an increase in two of the shedding rates, giving rise to four scenarios:

1. **Increased contribution from C:** A 50% increase in spc_C, meaning colonized individuals shed more spores when they contact a fomite, would increase ρ_{CH} from 0.057 to 0.085 and ρ_{CL} from 0.029 to 0.043.
2. **Increased contribution from D:** Likewise, a 50% increase in spc_D, meaning diseased individuals shed more spores when they contact a fomite, would increase ρ_{DH} from 0.123 to 0.188 and ρ_{DL} from 0.063 to 0.096.
3. **Increased contacts with P_H:** A 50% increase in cpd_H, meaning high-touch fomites are contacted more often, would increase ρ_{CH} from 0.057 to 0.085 and ρ_{DH} from 0.123 to 0.184. If high-touch fomites are contacted 50% more often, then $\omega = cpd_H/cpd_L$ would also increase by 50% from 1.96 to 2.94.
4. **Increased contacts with P_L:** Likewise, a 50% increase in cpd_L, meaning low-touch fomites are contacted more often, would increase ρ_{CL} from 0.029 to 0.043 and ρ_{DL} from 0.063 to 0.094. If low-touch fomites are contacted 50% more often, then $\omega = cpd_H/cpd_L$ would also decrease by 33.3% from 1.96 to 1.31.

The weighting constant for high-touch fomites, ω, is tied with the colonization rate upon transfer of spores from a fomite, β, as the most impactful parameter to the incidence of colonized due to high-touch fomites, as shown in Table 3.6, and has a positive significant

PRCC value for this outcome. A 50% increase in ω means that patients are contacting high-touch fomites more or contacting low-touch fomites less, giving rise to two more scenarios involving changes in the shedding rates:

5. **Increased contacts with P_H:** A 50% increase in $\omega = cpd_H/cpd_L$ from 1.96 to 2.94 due to more contacts with high-touch fomites would increase cpd_H by 50%, and cause ρ_{CH} to increase from 0.057 to 0.085 and ρ_{DH} to increase from 0.123 to 0.184. This scenario is identical to Scenario #3 above.

6. **Decreased contacts with P_L:** A 50% increase in $\omega = cpd_H/cpd_L$ from 1.96 to 2.94 due to less contacts with low-touch fomites would decrease cpd_L by 33.3%, and cause ρ_{CL} to decrease from 0.029 to 0.019 and ρ_{DL} to decrease from 0.063 to 0.041.

The results of the stochastic simulations of these six scenarios are given in Table 3.8. Note that Scenario #3 and #5 are identical situations.

Increased contacts with P_H, Scenario #3 & #5, sees the average incidence of colonized increase from 26.73 in the baseline scenario to 57.99 individuals, while individual trajectories within one standard deviation of the mean could have as many as 77.90 colonized individuals. This scenario produces the largest increase in the average and variability around that average of the incidence of colonized for all of the scenarios considered in this study. Unsurprisingly, increased contacts with P_L, Scenario #6, sees the largest increase in the average contribution of high-touch fomites incidence of colonized at 89% of colonizations compared with 77% in the reference scenario.

3.4.3 Results from Increasing the Efficacy of Cleaning and Disinfecting

Often, compliance with the varied disinfecting and cleaning protocols of hospital fomites is sub-optimal. Scenario #7 considers the impact of increasing the cleaning and disinfection parameters μ and σ even though neither had highly ranked significant PRCCs for any of the outcomes. The results of the stochastic simulation of Scenario #7 are given in Table 3.8 in which σ was increased from 0.83 to 1.0, the largest value in its range and μ was increased by 50% from 0.66 to 0.99.

Table 3.8: Comparison of the incidence results from the stochastic simulation of the system (3.1) with baseline parameter values from Table 3.2 (Scenario #0) to scenarios with increased values of parameters that the sensitivity analysis found to be influential on the incidence outcomes. Results are reported as the mean ± standard deviation of new cases per 1,000 admissions in 100 trajectories of the system. It can be assumed that the incidence results of the deterministic simulation of each scenario (not shown) are approximately equal to the mean incidence values of the stochastic simulation of each scenario.

	Scenario #0 (baseline values)	**Scenario #1** (more spores from C)	**Scenario #2** (more spores from D)
Incidence of C	26.73 ± 11.99	37.03 ± 15.24	36.26 ± 15.57
due to P_H	20.63 ± 10.20	28.25 ± 13.01	27.36 ± 12.93
due to P_L	6.09 ± 4.98	8.78 ± 6.06	8.90 ± 5.81
Incidence of D	21.35 ± 9.64	23.90 ± 10.29	22.48 ± 9.72

	Scenario #3 & #5 (more contacts with P_H)	**Scenario #4** (more contacts with P_L)	**Scenario #6** (less contacts with P_L)
Incidence of C	57.99 ± 19.91	24.27 ± 11.16	40.02 ± 16.63
due to P_H	49.95 ± 17.20	14.39 ± 8.29	35.57 ± 15.26
due to P_L	8.04 ± 6.34	9.88 ± 5.98	4.45 ± 4.57
Incidence of D	24.71 ± 9.25	19.89 ± 9.09	23.68 ± 10.98

	Scenario #7 (increased μ and σ)	**Scenario #8** (increased β)	**Scenario #9** (increased α)
Incidence of C	22.74 ± 10.65	43.27 ± 14.91	32.79 ± 15.83
due to P_H	17.29 ± 8.73	33.36 ± 12.25	25.91 ± 13.06
due to P_L	5.44 ± 4.84	9.91 ± 7.01	6.89 ± 6.21
Incidence of D	22.54 ± 9.13	24.62 ± 8.70	22.72 ± 10.46

	Scenario #10 (increased ϕ)	**Scenario #11** (increased k_R)
Incidence of C	28.25 ± 13.37	23.51 ± 10.78
due to P_H	21.09 ± 10.97	18.44 ± 9.45
due to P_L	7.16 ± 5.68	5.07 ± 3.85
Incidence of D	31.14 ± 11.13	20.19 ± 8.99

Since these two parameters do not have highly ranked significant PRCCs, it is not surprising that the results of the stochastic simulation of this scenario are similar to the results of the baseline scenario. This likely occurs since neither parameter directly affects the patient classes, but instead reduces the amount of spores in the environment. However, of all scenarios considered in this study, this scenario produced the smallest variation in the incidence of colonized. Also, this scenario was the only one to reduce the average ending pathogen level on high-touch fomites, from 0.09 spores per square centimeter in the baseline scenario to 0.07 spores per square centimeter. All other scenarios had an average pathogen level greater than 0.09 spores per square centimeter after 40 days. Interestingly, a similar result was not seen for low-touch fomites. The average ending pathogen level on low-touch fomites decreased from 0.05 to 0.04 spores per square centimeter, but other scenarios decreased to 0 spores per square centimeter.

3.4.4 Results from Changing Non-Environmental Rates

In addition to the scenarios that are strictly related to *C. difficile* transfer through the environment, changes in other disease progression rates were considered.

The rate of colonization upon transfer of spores, β, is in the top three most impactful parameters in all three outcomes and has positive significant PRCCs for all outcomes. Thus, for an increase in β, an increase in the outcome is expected. Scenario #8 considers when β is increased by 50% from 0.338 to 0.507 and results are given in Table 3.8. The average incidence of colonized increases from 26.73 to 43.27, while a smaller increase is seen in the incidence of diseased. Of all scenarios considered in this study, this scenario produced the smallest variation in the incidence of diseased.

The antibiotic prescription rate, α, is the third most impactful parameter to the incidence of colonized due to both high-touch and low-touch fomites and has positive significant PRCC values for both of these outcomes. Scenario #9 considers when α is increased by 50% from 0.5 to 0.75 and results are given in Table 3.8, where interestingly, the average contribution of high-touch fomites to new colonizations increases from 77% to 79%.

The disease rate of colonized individuals, ϕ, is the most impactful parameter to the incidence of diseased and has a positive significant PRCC value for this outcome. The

results of the stochastic simulation with ϕ increased by 50% from 0.024 to 0.036 are given in Scenario #10 in Table 3.8. In this scenario, the average incidence of diseased increases from 21.35 to 31.14 individuals, while individual trajectories within one standard deviation of the mean could have as many as 42.27 diseased individuals. Of all scenarios considered in this study, this scenario produces the largest increase in the average and variability around that average of the incidence of diseased. An interesting result of this scenario is that the average contribution of high-touch fomites to new colonizations decreases from 77% to 75% of colonizations.

Since disinfecting occurs when a patient is discharged, increasing the discharge rates of the patient classes would increase disinfecting and thus decrease the amount of pathogen present on fomites. This would in turn decrease the number of colonized patients and thus decrease the number of diseased patients. However, k_R is the only discharge rate that has a highly ranked significant PRCC value in any of the three outcomes. This is probably due to k_R being significantly larger than the other discharge rates and would therefore have the largest impact on the number of new admissions. The discharge rate of resistant patients, k_R, is the second most impactful parameter to the incidence of diseased and has a positive significant PRCC value for this outcome. The results of the stochastic simulation with k_R increased by 50% from 0.33 to 0.495 are given in Scenario #11 in Table 3.8, where the average incidence of diseased decreases from 21.35 to 20.19 individuals. When a discharge rate is increased, more patients are admitted into all four patient population classes. Since more patients will be admitted to the susceptible and colonized classes than to the diseased class according to the admission proportions, there will be more patients that will then move from these classes into the diseased class and be counted as newly diseased.

3.5 Future Work

The description of the transmission of *C. difficile* among patients through contacts with fomites is incomplete in our model. A major contributor to the transmission of *C. difficile* through environmental pathways is healthcare workers, which is not explicitly considered in this study. Healthcare workers directly contact patients and fomites, and transfer spores

between them. To include healthcare workers in this model, a class for healthcare workers would need to be added, as well as a description of how that class interacts with the patient and environmental reservoir classes and how colonization of patients occurs through those interactions.

It is important to note that, due to the formulation of the model, diseased individuals could not be linked with the fomite type that caused their infection. Reformulating the model to include two colonized, C_H and C_L, and two diseased, D_H and D_L, patient classes could allow for more exploration into the contribution of fomite touch frequency on incidence of diseased individuals.

The model in this study could be further improved by considering classifications of fomites beyond high-touch and low-touch fomites. The CDC's Guidelines for Environmental Infection Control in Healthcare Facilities [28] and for Disinfection and Sterilization in Healthcare Facilities [53] include recommended cleaning strategies for a variety of fomites. Further, the effect of contact time with a fomite on incidence could improve the description of transmission. However, more data would be needed on contacts with and transfer of spores from these different fomites to estimate parameter values.

As a comparison, the data set used to parameterize the model described in Lanzas et al. [37], included 11,046 admissions with 157 *C. difficile* infection cases total, leading to an incidence of new diseased of 14.21 per 1,000 admissions. Additionally, data obtained from the CDC's website for U.S. Acute Care Hospitals for 2018 and 2019 [14] and from the American Hospital Directory [5], yielded incidence ranges of $1.05 - 3.33$ (2018) and $0.1 - 2.94$ (2019) diseased cases per 1,000 admissions. Altogether, the data suggests a range of $0.1 - 14.21$ cases per 1,000 admissions and suggests that this current study may be overestimating the contribution of fomites to new cases. One potential reason is that the data set from [37] had high incidence values since *C. difficile* infection rates in 2008 were higher than current rates. Another possible cause for the overestimation is that the model in this study did not address the removal of spores due to contact or hand-washing, though a more likely reason is that the patient classes and environmental reservoirs were assumed to be homogeneously mixed, meaning every patient in the hospital ward contacted all spores in the environment at the same rate. This is less realistic as patients are much

more likely to interact with spores in their own room, where other patients are not likely to be, than anywhere else. Further consideration of environmental pathways of transmission should address spatial heterogeneity.

3.6 Conclusions

The difficulty of studying environmental transmission of healthcare-associated pathogens and a lack of understanding of these dynamics has hindered the ability to control *C. difficile* infections in healthcare settings. To gain a better understanding of the dynamics of *C. difficile* transmission, this study investigated the relative contribution of two environmental reservoirs to *C. difficile* transmission in a healthcare setting by explicitly including high-touch and low-touch fomites in a model based on the previous mathematical model by Lanzas et al. [37]. Additionally, this study aimed to determine the factors that influence the relative contribution from these two types of fomites. The treatment of environmental pathways in disease transmission in this study was novel in the distinction of types of fomites, including the tracking of bacterial spore populations on those fomites; the distinction between disinfecting and cleaning of the fomites; and the associated transmission, shedding, and cleaning parameters. An abundance of studies addressing patient-fomite contacts and transfer of pathogen allowed for this novelty.

Results of the sensitivity analysis determined scenarios to consider in simulations of the model. In most scenarios, on average $21 - 25\%$ of colonized cases were due to contact with low-touch fomites, while $75 - 79\%$ were due to contact with high-touch fomites, despite the extra cleaning of high-touch fomites. The two scenarios in which the ratio of contacts per day with high-touch to low-touch fomites, ω, was increased due to either an increase in the contacts per day with high-touch fomites (Scenario #3 and #5) or a decrease in the contacts per day with low-touch fomites (Scenario #6), had, on average, 86% and 89% of colonized cases due to a contact with high-touch fomites and 14% and 11% due to a contact with low-touch fomites, respectively. This suggests that when high-touch fomites are contacted much more often than low-touch fomites, high-touch fomites contribute much more to the incidence of colonized individuals. When high-touch fomites are not contacted much more

often than low-touch fomites, there is a smaller difference in the contribution from each fomite type to the incidence of colonized individuals.

Additionally, these two scenarios in which the ratio of contacts per day with high-touch to low-touch fomites, ω, was increased, were also the only two scenarios in which the average number of colonized patients was over 4% of the total admitted patients. This suggests that when high-touch fomites are responsible for a higher percentage of colonized cases, the colonization incidence will be larger.

Surprisingly, the disinfecting and cleaning parameters, σ and μ, were not identified as impactful parameters from the sensitivity analysis. Scenario #7 considered near perfect implementation of disinfecting and cleaning protocols. Since the average incidence of colonized and diseased individuals did not significantly decrease in this scenario, it implies that increasing the effectiveness of disinfecting and cleaning alone may not be sufficient to control an outbreak of *C. difficile* in a hospital ward. However, this was the only scenario to show a decrease in the ending pathogen levels on high-touch fomites. This result, combined with the results about high-touch fomites discussed above, suggests that a combined effort of increasing the effectiveness of disinfecting and cleaning of high-touch fomites and reducing contacts with high-touch fomites could reduce the incidence of colonized individuals in an outbreak.

Also impactful to the incidence of both colonized and diseased individuals was the rate at which spores were transferred from a fomite causing colonization, β. Manipulating the probability of colonization given the exposure dose would involve approaches that improve gut microbiota recovery.

The parameter that had the largest effect on the average percentage of patients who became diseased, as well as giving rise to the largest variation in the incidence of diseased individuals, was the parameter, ϕ, the rate at which colonized individuals with an insufficient immune response developed symptoms of *C. difficile* infection, as described in Scenario #10. Typically, hospitals do not test for *C. difficile* infection unless patients are exhibiting symptoms. Testing for asymptomatic individuals could help to reduce this rate, and thus reduce the incidence of diseased individuals, if testing is followed up by antibiotic stewardship practices. Overall, the factors that drive the progression from asymptomatic colonization

to infection are not well understood. However, antibiotic use has been identified as a risk factor for the progression to disease; therefore, antibiotic stewardship practices that target colonized patients may modify ϕ toward favorable results [10]. Vaccination with toxoid vaccines could also reduce the number of patients transitioning from colonized to diseased as vaccination may increase the proportion of individuals with an adequate immune response against *C. difficile* toxins.

Two additional modifiable parameters were also impactful to the incidence of colonized and diseased individuals, respectively, the antibiotic prescription rate, α, and the discharge rate of resistant individuals, k_R. An increase in α in Scenario #9 produced an increase in the incidence of colonized individuals since more prescriptions for antibiotics produces more susceptible individuals. An increase in k_R in Scenario #11 produced a decrease in the incidence of diseased individuals. These results indicate that a decreased duration of a hospital visit for a resistant individual (e.g., increased k_R) leads to a decreased chance of being prescribed an antibiotic and becoming susceptible to a *C. difficile* infection. Patients who stay in the hospital for longer periods of time are likely in the hospital for more serious conditions and may be more likely to be prescribed an antibiotic. Thus, reducing the rate at which antibiotics are prescribed and reducing the duration of patient visits should reduce incidence of colonized and diseased individuals. This concurs with results from previous modeling work [8].

The results of this study suggest potential disease control strategies: the combined effort of increasing the effectiveness of disinfecting upon discharge and extra regular cleaning of high-touch fomites and reducing patient contacts with high-touch fomites; testing and monitoring asymptomatic patients; reducing the rate at which antibiotics are prescribed; and reducing the duration of patient stays in the hospital. By explicitly considering the environmental pathways of *C. difficile* transmission, more emphasis can be placed on the effectiveness of these control strategies leading to a decrease in the incidence of colonized and diseased individuals. Ultimately, this study has demonstrated the need to further study the role of environmental pathways in the transmission of *C. difficile* in a healthcare setting.

Bibliography

[1] Adelman, D., Murray, J., Wu, T.-T., Mäki, M., Green, P., and Kelly, C. (2018). Measuring change in small intestinal histology in patients with celiac disease. *Am J Gastroenterol*, 113:339–347. 7, 8, 9, 10, 12, 14, 15, 16, 17, 18, 19, 37, 39, 40

[2] Ahn, T.-H., Cao, Y., and Watson, L. (2008). Stochastic simulation algorithms for chemical reactions. In *BIOCOMP*. 93, 94

[3] Ali, S., Muzslay, M., and Wilson, P. (2015). A novel quantitative sampling technique for detection and monitoring of *Clostridium difficile* contamination in the clinical environment. *Journal of Clinical Microbiology*, 53(8):2570–2574. 81

[4] Allen, L. (2010). *Stochastic Population and Epidemic Models: Persistence and Extinction*. Springer. 93

[5] American Hospital Directory (accessed December 14, 2020). *Hospital Statistics by State*. Louisville, KY. 110

[6] Argoti, A., Fan, L., Cruz, J., and Chou, S. (2008). Introducing stochastic simulation of chemical reactions using the gillespie algorithm and matlab: Revisited and augmented. *Chemical Engineering Education*, 42:35–46. 93, 94

[7] Barthel, E. R. (2017). On the utility of a compartmental population kinetics model of intestinal epithelial stem cell proliferation and differentiation. *Theor Biol Med Model*, 14:25. 12, 34, 35, 36

[8] Bintz, J., Lenhart, S., and Lanzas, C. (2017). Antimicrobial stewardship and environmental decontamination for the control of *Clostridium difficile* transmission in healthcare settings. *Bulletin of Mathematical Biology*, 79(1):36–62. 82, 113

[9] Carling, P. C., von Beheren, S., Kim, P., and Woods, C. (2008). Intensive care unit environmental cleaning: An evaluation in sixteen hospitals using a novel assessment tool. *Journal of Hospital Infection*, 68:39–44. 81

[10] Caroff, D. A., Yokoe, D. S., and Klompas, M. (2017). Evolving insights into the epidemiology and control of *Clostridium difficile* in hospitals. *Clinical Infectious Diseases*, 65(7):1232–1238. 113

[11] Celiac Disease Foundation (2016 (accessed March 2, 2021)). *9 Questions About Celiac Disease, Answered, as Featured in USA Today*. 8, 9, 13, 26

[12] Cells Applications (accessed 2020-05-05). *Human Intestinal Epithelial Cells: HInEpC*. https://www.cellapplications.com/human-intestinal-epithelial-cells-hinepc. 10, 12

[13] Centers for Disease Control and Prevention (2019). Antibiotic resistance threats in the United States. Technical report, U.S. Department of Health and Human Services, Atlanta, GA. 80

[14] Centers for Disease Control and Prevention (accessed December 14, 2020). *Current HAI Progress Report*. Atlanta, GA. 110

[15] Chang, H. T., Krezolek, D., Johnson, S., Parada, J. P., Evans, C. T., and Gerding, D. N. (2007). Onset of symptoms and time to diagnosis of *Clostridium difficile*-associated disease following discharge from an acute care hospital. *Infection Control & Hospital Epidemiology*, 28(8). 88, 89

[16] Chen, S., Lenhart, S., Day, J., Lee, C., Dulin, M., and Lanzas, C. (2018). Pathogen transfer through environment-host contact: An agent-based queueing theoretic framework. *Mathematical Medicine, and Biology*, 35:409425. 82

[17] Cheng, V. C. C., Chau, P. H., Lee, W. M., Ho, S. K. Y., Lee, D. W. Y., So, S. Y. C., Wong, S. C. Y., Tai, J. W. M., and Yuen, K. Y. (2015). Hand-touch contact assessment of high-touch and mutual-touch surfaces among healthcare workers, patients, and visitors. *Journal of Hospital Infection*, 90:220–225. 88, 89, 92

[18] Cohen, I. S., Day, A., and Shaoul, R. (2019). Gluten in celiac disease - more or less? *Rambam Maimonides Med J*, 10. 9

[19] Cooper, J. (2005). The poisson and exponential distributions. *Applied Probability Trust*, page 123. 93

[20] Crosnier, C., Stamataki, D., and Lewis, J. (2006). Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nature Reviews Genetics*, 7:349–359. 12, 13, 34, 36

[21] Demin, O., Smirnov, S., Sokolov, V., Cucurull-Sanchez, L., Pichardo-Almarza, C., Flores, M., Benson, N., and Demin, O. V. (2013). Modeling of celiac disease immune response and the therapeutic effect of potential drugs. *BMC Systems Biology*, 7:56. 7, 19, 20, 34, 36, 74

[22] Dickson, B., Streutker, C., and Chetty, R. (2006). Coeliac disease: an update for pathologists. *J Clin Pathol*, 59:1008–1016. 7, 10, 12, 34, 35

[23] Dubberke, E. R., Reske, K. A., Noble-Wang, J., Thompson, A., Killgore, G., Mayfield, J., Camins, B., Woeltje, K., McDonald, J. R., McDonald, L. C., and Fraser, V. J. (2007). Prevalence of *Clostridium difficile* environmental contamination and strain variability in multiple health care facilities. *American Journal of Infection Control*, 35(5):315–318. 81

[24] Forde, J. and Meeker, B. (2009). A mathematical model of celiac disease. Unpublished. 19

[25] Fox, L. (2020). *Two Applications of Mathematical Modeling: Control Mechanisms of Heart Rate Variability and Contribution of Environmental Pathways in Disease Transmission*. PhD thesis, The University of Tennessee, Knoxville. 94

[26] Gerding, D. N., Johnson, S., Peterson, L. R., Mulligan, M. E., and Silva, J. (1995). *Clostridium difficile*-associated diarrhea and colitis. *Infection Control and Hospital Epidemiology*, 16:459–477. 81

[27] Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25). 3, 83, 93, 94

[28] Healthcare Infection Control Practices Advisory Committee (2003). *Guidelines for Environmental Infection Control in Health-Care Facilities*. Centers for Disease Control and Prevention. Updated 2019. 81, 110

[29] Holling, C. S. (1959). Some characteristics of simple types of predation and parasitism. *The Canadian Entomologist*, 91:385–398. 92

[30] Huang, Q., Horn, M. A., and Ruan, S. (2019). Modeling the effect of antibiotic exposure on the transmission of methicillin-resistant *Staphylococcus aureus* in hospitals with environmental contamination. *Mathematical Biosciences and Engineering*, 16(5):3641–3673. 82

[31] Huslage, K., Rutala, W. A., Sickbert-Bennett, E., and Weber, D. (2010). A quantitative approach to defining "high-touch" surface in hospitals. *Infection Control and Hospital Epidemiology*, 31(8):850–853. 3, 82

[32] Kawabata, S., Boyaka, P., Coste, M., Fujihashi, K., Yamamoto, M., McGhee, J., and Kiyono, H. (1998). Intraepithelial lymphocytes from villus tip and crypt portions of the murine small intestine show distinct characteristics. *Gastroenterology*, 115:866–873. 13, 15

[33] Kaye, R. and Konz, S. (1986). Volume and surface area of the hand. *Proceedings of the Human Factors Society*. 88, 89

[34] Kramer, A., Schwebke, I., and Kampf, G. (2006). How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. *BMC Infectious Diseases*, 6(130):1–8. 81

[35] Kyne, L., Warny, M., Qamar, A., and Kelly, C. P. (2000). Asymptomatic carriage of *Clostridium difficile* and serum levels of IgG antibody against toxin A. *New England Journal of Medicine*, 342(6):390–397. 89

[36] Lähdeaho, M.-L., Mäki, M., Laurila, K., Huhtala, H., and Kaukinen, K. (2011). Small-bowel mucosal changes and antibody responses after low- and moderate-dose gluten challenge in celiac disease. *BMC Gastroenterology*, 11:129. 10, 22, 65, 66, 67, 68, 69, 71, 72, 73, 75, 76, 78

[37] Lanzas, C., Dubberke, E. R., Lu, Z., Reske, K. A., and Gròhn, Y. T. (2011). Epidemiological model for *Clostridium difficile* transmission in healthcare settings.

Infection Control and Hospital Epidemiology, 32(6):553–561. 83, 85, 87, 88, 97, 110, 111

[38] Lawley, T. D., Clare, S., Deakin, L. J., Goulding, D., Yen, J. L., Raisen, C., Brandt, C., Lovell, J., Cooke, F., Clark, T. G., and Dougan, G. (2010). Use of purified *Clostridium difficile* spores to facilitate evaluation of health care disinfection regimens. *Applied and Environmental Microbiology*, 76(20):6895–6900. 88, 92

[39] Lawson, P. A., Citron, D. M., Tyrrell, K. L., and Finegold, S. M. (2016). Reclassification of *Clostridium difficile* and *Clostridioides difficile* (Hall and O'Toole 1935) Prévot 1938. *Anaerobe*, 40:95–99. 80

[40] Leffler, D. A. and Lamont, J. T. (2015). *Clostridium difficile* infection. *The New England Journal of Medicine*, 372(16):1539–1548. 81

[41] Lessa, F. C., Mu, Y., Bamberg, W. M., Beldavs, Z. G., Dumyati, G. K., Dunn, J. R., Farley, M. M., Holzbauer, S. M., Meek, J. I., Phipps, E. C., Wilson, L. E., Winston, L. G., Cohen, J. A., Limbago, B. M., Fridkin, S. K., Gerding, D. N., and McDonald, L. C. (2015). Burden of *Clostridium difficile* infection in the United States. *The New England Journal of Medicine*, 372(9):825–834. 2, 80

[42] Lukes, D. (1982). *Differential Equations: Classical to Controlled*. Academic Press. 27

[43] MacLeod-Glover, N. and Sadowski, C. (2010). Efficacy of cleaning products for *C. difficile*. *Canadian Family Physician*, 56:417–423. 81

[44] Marino, S., Hogue, I., Ray, C., and Kirschner, D. (2008). A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology*, 254:178–196. 96, 97, 99

[45] MathWorks (2019). MATLAB, R2019b. Computer Software. 38, 99

[46] McFarland, L. V. (2008). Update on the changing epidemiology of *Clostridium difficile*-associated disease. *Nature Clinical Practice Gastroenterology and Hepatology*, 5(1):40–48. 88, 89

[47] Orenstein, R., Aronhalt, K. C., McManus, J. E., J., and Fedraw, L. A. (2011). A targeted strategy to wipe out *Clostridium difficile*. *Infection Control and Hospital Epidemiology*, 32(11):1137–1139. 81

[48] Otter, J. A., Yezli, S., and French, G. L. (2011). The role played by contaminated surfaces in the transmission of nosocomial pathogens. *Infection Control and Hospital Epidemiology*, 32(7):687–699. 81

[49] Otter, J. A., Yezli, S., Salkeld, J. A. G., and French, G. L. (2013). Evidence that contaminated surfaces contribute to the transmission of hospital pathogens and an overview of strategies to address contaminated surfaces in hospital setting. *American Journal of Infection Control*, 41:S6–S11. 81

[50] Parker, A., Maclarn, O., Fletcher, A., Muraro, D., Kreuzaler, P., Byrne, H., Maini, P., Watson, A., and Pin, C. (2017). Cell proliferation within small intestinal crypts is the principle driving force for cell migration on villi. *FASEB*, 31. 19, 20, 21, 24, 34, 35, 36

[51] Rafii, F., Sutherland, J. B., and Cerniglia, C. E. (2008). Effects of treatment with antimicrobial agents on the human colonic microflora. *Therapeutics and Clinical Risk Management*, 4(6):1343–1357. 88, 89

[52] Rubin, M. A., Jones, M., Leecaster, M., Khader, K., Ray, W., Huttner, A., Huttner, B., Toth, D., Sablay, T., Borotkanics, R. J., Gerding, D. N., and Samore, M. H. (2013). A simulation-based assessment of strategies to control *Clostridium difficile* transmission and infection. *PLoS ONE*, 8(11):1–11. 82

[53] Rutala, W. A., Weber, D. J., and Healthcare Infection Control Practices Advisory Committee (2008). *Guideline for Disinfection and Sterilization in Healthcare Facilities, 2008*. Centers for Disease Control and Prevention. 81, 110

[54] Sethi, A. K., Al-Nassir, W. N., Nerandzic, M. M., Bobulsky, G. S., and Donskey, C. J. (2010). Persistence of skin contamination and environmental shedding of *Clostridium difficile* during and after treatment of *C. difficile* infection. *Infection Control and Hospital Epidemiology*, 31(1):21–27. 88, 89

[55] Shams, A. M., Rose, L. J., Edwards, J. R., Cali, S. Harris, A. D., Jacob, J. T., LaFae, A., Pineles, L. L., Thom, K. A., McDonald, L. C., Arduino, M. J., and Noble-Wang, J. A. (2016). Assessment of the overall and multidrug-resistant organism bioburden on environmental surfaces in healthcare facilities. *Infection Control and Hospital Epidemiology*, pages 1–7. 81, 88, 91

[56] Shaughnessy, M. K., Micielli, R. L., DePestel, D. D., Arndt, J., Strachan, C. L., Welch, K. B., and Chenoweth, C. E. (2011). Evaluation of hospital room assignment and acquisition of *Clostridium difficile* infection. *Infection Control and Hospital Epidemiology*, 32(3):201–206. 81

[57] Smith, H. (1995). Monotone dynamical systems: An introduction to the theory of competitive and cooperative systems. *Mathematical Surveys and Monographs*, 41. 27

[58] Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H., and Kuijper, E. J. (2016). *Clostridium difficile* infection. *Nature Reviews: Disease Primers*, 2:16020. 2, 80, 81

[59] Sompayrac, L. (2015). *How the Immune System Works*. Wiley-Blackwell. 6, 7

[60] Starr, J. M., Campbell, A., Renshaw, E., Poxton, I. R., and Gibson, G. J. (2009). Spatio-temporal stochastic modelling of *Clostridium difficile*. *Journal of Hospital Infection*, 71:49–56. 82

[61] Stephenson, B., Lanzas, C., Lenhart, S., and Day, J. (2017). Optimal control of vaccination rate in an epidemiological model of *Clostridium difficile* transmission. *Journal of Mathematical Biology*, 75(6-7):1693–1713. 87, 88

[62] Stephenson, B., Lanzas, C., Lenhart, S., Ponce, E., Bintz, J., Dubberke, E. R., and Day, J. (2020). Comparing intervention strategies for reducing *Clostridioides difficile* transmission in acute healthcare settings: An agent-based modeling study. *BMC Infectious Diseases*. 82

[63] Trask, J. M. (2014). Modeling celiac disease. Master's thesis, The University of Tennessee, Knoxville. 19

[64] United States Food and Drug Administration (2020 (accessed March 2, 2021)). *Gluten-Free Labeling of Foods*. 9

[65] Verity, P., Wilcox, M. H., Fawley, W., and Parnell, P. (2001). Prospective evaluation of environmental contamination by *Clostridium difficile* in isolation side rooms. *Journal of Hospital Infection*, 49:204–209. 81

[66] Weber, D. J., Anderson, D. J., Sexton, D. J., and Rutala, W. A. (2013). Role of the environment in the transmission of *Clostridium difficile* in health care facilities. *American Journal of Infection Control*, 41:S105–S110. 81

[67] Wolkewitz, M., Dettenkofer, M., Bertz, H., Schumacher, M., and Huebner, J. (2008). Environmental contamination as an important route for the transmission of the hospital pathogen VRE: Modeling and prediction of classical interventions. *Infectious Disease: Research and Treatment*, 1:3–11. 82

Appendix

A Abbreviations

Cd - crypt depth

CeD - celiac disease

CI - confidence interval

CVEU - crypt-villus epithelial unit

GFD - gluten-free diet

HLA - human leucocyte antigen

IEL - intraepithelial lymphocyte, also known as natural killer (NK) cells

TG-2 - tissue transglutaminase-2

Vh - villus height

Vh:Cd - villus height-to-crypt depth ratio

Vita

Cara Jill Sulyok was born in Lansdale, Pennsylvania and raised by her parents, Gabriel D. and Clarisse Sulyok. She graduated from Lansdale Catholic High School in 2011 and began her undergraduate career at Ursinus College in Collegeville, Pennsylvania that fall. Cara always knew she wanted to pursue a career in mathematics education, but quickly uncovered her passion for research at the intersection of mathematics and biology while working on her undergraduate honors thesis. Under the advisement of Mohammed Yahdi, her thesis utilized ordinary differential equations and numerical simulations to provide a mathematical framework for designing cost-effective and environmentally safe strategies to minimize plant damage in agroecosystems. She graduated from Ursinus College with her Bachelor of Science in Mathematics and minors in Statistics and Secondary Education in May 2015.

In August 2015, Cara began her graduate career at the University of Tennessee, Knoxville (UTK), studying mathematical biology and numerical mathematics. She was afforded the opportunity to teach multiple classes as well as work with Malissa Meadows to develop course materials for coordinated courses. Advised by Judy Day and Suzanne Lenhart, her graduate work involved formulating two mathematical models - one to explore the immune response in those with celiac disease and one to quantify the contribution of environmental pathways to *Clostridioides difficile* transmission in healthcare settings. As a patient with celiac disease, Cara hopes to find a viable treatment other than the gluten-free diet using mathematics.

Cara graduated from UTK in May 2021 with a Doctor of Philosophy in Mathematics. In August 2021, Cara will join Lewis University in Romeoville, Illinois as an Assistant Professor of Mathematics. She looks forward to helping her students develop and deepen their passion for mathematics, guiding her students in interdisciplinary research projects, and continuing to improve and grow as an educator.