Boston College

William F. Connell School of Nursing

IDENTIFICATION OF OPTIMAL STOOL DONOR HEALTH AND INTESTINAL MICROBIOME CHARACTERISTICS FOR FECAL MICROBIOTA TRANSPLANTATION

a dissertation

by

NANCY E. DUBOIS, MSN, MBA, RN

submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

May 2019

Identification of Optimal Stool Donor Health and Intestinal Microbiome Characteristics for Fecal Microbiota Transplantation

Nancy E. Dubois

Chair: Catherine Y. Read, PhD, RN

Abstract

Background. *Clostridium difficile* infections (CDI) account for 20-30% of healthcare-acquired infections, resulting in serious patient and economic burdens. CDI incidence has grown rapidly due to overuse of antibiotics and an aging population, posing a significant public health threat. Fecal microbiota transplantation (FMT) using donor stool has demonstrated clinical efficacy rates up to 94% and long-term restoration of a healthy intestinal microbiome. Challenges with donor screening, lack of research about optimal stool donor characteristics and intestinal microbiome composition, and a poorly fit screening model, create barriers to the availability of FMT.

Purpose. This study aimed to generate essential information about FMT donor characteristics predictive of passing the screening and donor intestinal microbiome compositions associated with FMT clinical efficacy. The primary aims were to 1) identify previously unstudied characteristics of prospective FMT donors that are predictive of passing a stool bank's screening process; and 2) determine whether donor intestinal microbial diversity is related to FMT clinical efficacy in preventing recurrent CDI.

Methods. This study was conducted as a secondary analysis on a cohort of previously screened donors (n=770). Aim 1 was tested through a logistic regression of donor characteristics (gender, age, body mass index, frequency of bowel movements, diet, tobacco and alcohol use, and

seasonality) with screening outcomes. Aim 2 was tested through a simple regression evaluating donor intestinal microbial diversity and rates of FMT clinical efficacy.

Results. One donor characteristic in the logistic regression, frequency of bowel movements (p = 0.018), was significantly predictive of whether a donor passed the screening. Specifically, donors who had fewer than two bowel movements per day were more likely to pass. All other characteristics were not predictive. Similarly, the linear regression evaluating alpha diversity and FMT clinical efficacy was not significantly predictive of clinical efficacy (p = 0.140).

Conclusion. Findings were used to support recommendations for improving prospective donor screening that nurses and other clinicians can implement to decrease challenging logistics, reduce costs and barriers, and potentially increase FMT clinical efficacy.

Acknowledgements

This dissertation would not have been possible without the extensive support from my committee and fellow classmates who offered critical and timely advice along the way. Dr. Catherine Read's commitment and invaluable guidance have been a core influence in seeing this dissertation come to fruition. Dr. Zain Kassam's belief in my capabilities and remarkable insights demonstrate his passion for this emerging area of medicine and the patients that it treats. His foresight was instrumental in bringing the ideas and concepts behind this dissertation together. Lastly, a heartfelt thank you must be given to Dr. Matt Gregas and Dr. Susan Kelly-Weeder who rallied quickly and stepped in when their expertise was critically needed.

This dissertation is the result of support from a long list of individuals, groups, and companies. Firstly, this accomplishment could not have taken place without the valuable financial support from Boston College and the International Society of Nurses in Genetics who graciously selected my project for funding. OpenBiome's mission to get this life-saving treatment into the hands of clinicians made this cohort evaluation possible. Most importantly are the many patients who placed their hopes on this investigational treatment, when for many of them there were no other options left.

Finally, this dissertation is dedicated to my family and dear friends, but most importantly, to my children Leo and Ellie, whose patience and understanding these many, many years have been remarkable. To Leo and Ellie – you both have been a constant beacon of light in this long journey. Know in your hearts that you can reach for you dreams no matter how difficult they may seem. "Never give in, never give in, never, never, never, never – in nothing, great or small, large or petty – never give in, except to convictions of honour and good sense." (Winston Churchill, Harrow School, October 29, 1941).

Table of Contents

List of Tables

List of Figures

Chapter 1. Statement of the Problem

Clostridium difficile infections (CDI) are a global concern and the primary cause of diarrhea and colitis in industrialized countries. Over the past decade, the incidence of CDI has increased significantly because of factors that reduce the number and diversity of commensal intestinal bacteria that normally keep *Clostridium difficile* in check. The overuse of antibiotics that alter the normal intestinal microbiome composition is a well-known contributor to CDI, and the problem is compounded by the demographic shift toward older people, whose intestinal microbiomes are less diverse than in younger people (Dubberke & Olsen, 2012; Lessa, Winston, & McDonald, 2015; McNabb-Baltar, Yaghoobi, O'Byrne, Soulellis, & Trinh, 2013).

Fecal microbiota transplantation (FMT) utilizing human stool donors has emerged as a viable treatment option for recurrent CDI, demonstrating clinical cure in up to 94% of recipients (Kassam, Lee, Yuan, & Hunt, 2013; van Nood, Vrieze, et al., 2013) and long-term restoration of a recipient's commensal bacteria and intestinal microbiome composition (Broecker et al., 2013). Fecal transplant material from healthy donors with diverse intestinal microbiomes effectively inhibits CDI through the recolonization of a patient recipient's intestines with commensal bacteria, thus preventing *Clostridium difficile* overgrowth and restoring patient health. Despite the documented success of FMT, the inability of nurses and other clinicians to expeditiously evaluate prospective donors to procure fecal material from healthy donors with consistently high rates of clinical efficacy continues to be a major barrier to patient access. As the demand for FMT treatments has increased, so has the need to generate knowledge about optimal prospective FMT donor characteristics and ideal intestinal microbiome compositions.

Information about FMT donor characteristics such as gender, diet, and body mass index is collected by nurses under the donor screening program currently utilized by the largest

international public stool bank located in the United States (U.S.), OpenBiome in Somerville, Massachusetts. Despite its availability, this information has not been adequately evaluated for its potential impact on the donor screening process, which may contribute to the notably low donor acceptance rate of 4%-10% encountered under existing screening programs (L. J. Burns et al., 2015; Paramsothy et al., 2015). Furthermore, the relationship between the level of bacterial diversity of the donor's intestinal microbiome (as measured by 16S ribosomal ribonucleic acid [rRNA] sequencing) that is transplanted during the FMT and clinical efficacy has not been examined in sufficiently robust stool donor cohorts. This gap may explain why the clinical efficacy of FMT, although high (van Nood, Vrieze, et al., 2013; Youngster et al., 2014), falls below 100%. Robust donor screening guidelines that address currently underutilized donor characteristics and information about the 16S rRNA level of bacterial diversity of the donor stool microbiome have the potential to improve and expedite the FMT donor screening process and increase clinical efficacy. These advances could contribute to the goal of reducing patient suffering and threats to overall public health posed by CDI.

The Intestinal Microbiome and *Clostridium difficile* Infections

The collective bacteria that live in the human intestines are known as the intestinal microbiome. Bacteria that comprise the intestinal microbiome in healthy individuals use various strategies to compete for scarce energy resources and maintain a balance in their population levels to prevent dominant colonization by any one bacterial species, a phenomenon known as colonization resistance. Colonization resistance is an effective protective mechanism that prevents overpopulation of disease-causing opportunistic bacteria such as *Clostridium difficile*. However, when perturbations to the intestinal microbiome occur, as when antibiotics eliminate critical normal flora, colonization resistance fails and pathogenic bacteria such as *Clostridium*

difficile can thrive and cause disease. Restoring colonization resistance through FMT can be an effective strategy for combatting intestinal diseases caused by opportunistic bacteria.

Fecal Microbiota Transplant as a Potentially Effective Solution

Fecal microbiota transplantation is an emerging medical intervention where fecal material from healthy donors is collected, minimally processed, and then transplanted into patients with recurrent CDI who have significantly suppressed intestinal bacteria. Clinically, fecal preparations are often infused by physicians either directly into the large intestine via colonoscopy or to the small intestine through upper endoscopy. With the recent availability of capsules, fecal preparations may now be dispensed for oral delivery under nursing or clinician supervision (Samuel, Crumb, & Duba, 2014).

Although the exact mechanism for how FMT prevents CDI recurrence is not entirely understood, transplantation of fecal material from healthy donors is believed to restore the intestinal microbiome by preemptively repopulating a CDI patient's intestines with diverse commensal bacteria (Broecker et al., 2013). The bacteria introduced into the patient's intestines via the donor fecal material restore colonization resistance and collectively out-compete and exclude pathogenic *Clostridium difficile* bacteria, thus preventing re-colonization, imbalances in the microbiota (called dysbiosis), and disease recurrence. Although the procedure is not new to medicine, recent randomized controlled clinical trials have demonstrated resolution rates of recurrent CDI in upwards of 94% of patients after one or two FMT treatments utilizing healthy donor derived fecal material (van Nood, Vrieze, et al., 2013; Youngster et al., 2014).

Statement of the Problem

The recent emergence of a hypervirulent *Clostridium difficile* strain BI/NAP1/027, an aging population at greater risk for intestinal dysbiosis and medical interventions, and continued

overuse of antibiotics have led to a rapid increase in demand for safe and efficacious donor fecal material to treat recurrent CDI (Borgia, Maraolo, Foggia, Buonomo, & Gentile, 2015; Khoruts & Weingarden, 2014). Despite the clinical efficacy of FMT supported within the literature and a permissive regulatory environment (Center for Biologics Evaluation and Research, 2013), significant barriers to access continue to exist. Impediments to patient access consistently reported by nurses and other clinicians are the procedures and costs associated with finding and screening suitable donors (Paramsothy et al., 2015). Although basic guidelines for evaluating donors for infectious and communicable diseases exist, pertinent donor characteristics purported in the literature to directly influence an FMT donor's health and potentially FMT clinical efficacy such as gender, diet, age, and microbiome diversity, have not been well explored and thus are not currently evaluated during the prospective FMT screening process. The dearth of information, largely precipitated by the lack of sufficiently robust donor cohorts, has created a notable patient barrier to safe and efficacious donor fecal material, FMT treatment, and CDI cure.

Significance of the Problem

Prior to the emergence of the first public stool bank in 2013 in the U.S., FMT donor screenings were solely conducted on a one-off basis by nurses and clinicians using a directed donor approach with the focus on minimizing the risk of transmitting infectious or communicable diseases. Under directed donations, prospective FMT donors are sourced for a one-time treatment donation from within the patient's friends and family network. In the absence of consensus guidelines about relevant donor information, prospective donor screening is non-cohesive and differs notably among clinicians and across institutions. Coupled with clinician and patient concerns about the distasteful aspects of the procedure and fear associated with handling

fecal material, institutional adoption of FMT has been slow. For these reasons, it has been difficult to generate the data needed to make inferences about donor characteristics that would support robust guidelines and reduce screening costs.

With the introduction of the first public stool bank, OpenBiome, the number of donors screened has increased dramatically. Despite this increase in donor data, few improvements have been made to the current conservative and logistically challenging donor screening approach, which continues to be time intensive, conducted over multiple in-person nursing and other clinician visits, and requires extensive laboratory screening for infectious and transmissible diseases (Bakken et al., 2011). The process, which takes several weeks to complete, has a reported 4-10% donor pass rate and deferral rates between 90-96% (L. J. Burns et al., 2015; Paramsothy et al., 2015). Additionally, despite the increasing access and affordability of microbiome-based genetic sequencing technologies, there is a paucity of current research on the intestinal microbiome and the characteristics of the microbiome of transplanted fecal material that may predict clinical efficacy of FMT. Thus, evidence to support the clinical utility of incorporating information about the donor intestinal microbiome diversity using 16S rRNA sequencing analysis during the prospective FMT donor screening process is lacking. Due to this lack of exploration of pertinent donor characteristics, such as age, gender, and intestinal microbiome, the prospective FMT donor screening approach has not evolved despite indications that such knowledge could lead to program tailoring, efficient donor targeting and recruitment, and reductions in program costs and treatment delays.

The current demanding and time-consuming approach to FMT donor screening has a negative impact on patient psychosocial well-being and quality of life (Pakyz, Moczygemba, VanderWielen, & Edmond, 2016). Patients who choose to use a directed donor approach take on

the entire burden of finding a suitable donor. However, under the current approach, ten or more prospective donors on average need to be identified by patients and subsequently screened by nurses and other clinicians in order to find one that is suitable (Paramsothy et al., 2015). This presents practical social and ethical challenges to patients who must disclose their private health information to individuals who might not otherwise be aware of the patient's health status. Similarly, FMT donors face social and ethical dilemmas from having to disclose their private health and lifestyle information. Anxiety resulting from the discovery of asymptomatic communicable diseases or from needing to provide a fecal sample is common among FMT donors (Paramsothy et al., 2015). Under these circumstances, finding a suitable family member, friend, or acquaintance to donate fecal material can be challenging and have a profound impact on CDI patients, requiring more counseling and other time-intensive nursing interventions.

Once a prospective donor has volunteered, out-of-pocket patient costs associated with donor testing and FMT treatment can increase rapidly, creating significant health disparities (Pakyz et al., 2016). Laboratory costs alone may surpass $500 per screening and are not reimbursable (Helwick, 2015). In cases where pre-screened, anonymous FMT donors from public stool banks are used, the patient's effort to find a suitable donor is reduced. However, while public stool banks attempt to minimize the costs of donor recruitment and fecal procurement, they encounter the same logistical challenges. Because the costs of finding and monitoring FMT donors are absorbed within the operational costs of the public stool bank, any inefficiency in donor screening directly impacts the prices public stool banks charge for anonymous fecal material. Currently there are no dedicated billing codes to cover fecal material from public stool banks or the FMT procedure. As such, patient out of pocket costs for one FMT treatment may cost upwards of $10,000 per treatment (Petrof & Khoruts, 2014). This further

shifts the burden of living with CDI disease to socially disadvantaged patients who lack sufficient economic resources to access FMT as a treatment for CDI.

Although patients suffering from recurrent CDI tend to be open and receptive to FMT, on average, patients wait approximately 7.2 months after diagnosis before receiving an FMT (Pakyz et al., 2016). This delay places a significant burden on both the healthcare system and the patients who undergo multiple and sometimes continuous long-term antibiotic treatment to avoid disease recurrence and mitigate disease progression. The overall cost of treating CDI with vancomycin for ten days, the current standard of care, is estimated to be $12,306 per recurrence and increases with each recurrence, length of antibiotic treatment, and CDI severity (Stranges, Hutton, & Collins, 2013). Recurrent CDI also negatively impacts a patient's pre-existing comorbidities and increases the complexity of disease-appropriate healthcare delivery. As such, approximately 29,000 deaths in the U.S. each year are attributed to complications arising from CDI infection (Lessa et al., 2015). Further, recurrent CDI has been shown to increase length of stay by 2.3 days on average and contribute to an additional 30-60% increase in healthcare costs (Dubberke & Olsen, 2012). Delays in access to FMT and CDI cure have a direct and cumulative financial and health impact on patients.

Most notably, the length of time that a patient suffers from CDI poses a significant public health threat. When a patient is diagnosed with CDI, the healthcare facility, nurses, and other clinicians must initiate stringent isolation and contact precautions, as well as enhanced infectious disease control measures to reduce contamination and carriage of *Clostridium difficile* spores that could lead to a CDI outbreak. While patients with CDI invariably spend some time in a hospital setting for treatment, much of CDI treatment has shifted to outpatient treatment in primary care settings and within the community. The incidence of community-acquired CDI is now estimated

to represent 20-40% of all CDI cases (Gupta & Khanna, 2014; Khanna et al., 2012). While hospital-acquired CDI tends to affect older patients with significant pre-existing comorbidities, the population affected by community-acquired CDI tends to be younger and have lower comorbidity scores (Khanna et al., 2012). The ability of CDI to effectively transition over the past decade into a new setting and population previously thought to be protected against CDI illustrates the challenges that nurses and other clinicians face in combating this robust communicable disease in the community and healthcare settings.

Delays in receiving FMT arising from challenges associated with sourcing suitable donors, a lack of safe donor fecal material availability, and high out of pocket medical costs lead to significant health and economic burdens for patients, the healthcare industry, and overall public welfare. The identification of optimal FMT donor characteristics could provide the information needed to improve the prospective FMT donor screening approach and allow for better targeting of prospective donors by nurses and other clinicians, thus significantly increasing the pass rate and reducing costs associated with recruitment and screening. Knowledge about donor 16S rRNA intestinal microbiome bacterial diversity and its relationship with FMT clinical efficacy could support the incorporation of intestinal microbiome sequencing as a biomarker during prospective FMT donor screening. In turn, the adoption of this technology could help to reduce out of pocket screening costs, shorten laboratory testing delays, and increase overall clinical efficacy of FMT. Combined, these improvements in the recruitment and screening of prospective FMT donors could reduce overall costs and barriers to institutional adoption, expedite patient access to FMT treatments, and ultimately reduce the incidence of CDI.

Purpose of the Study

The first purpose of this study was to identify additional, previously unstudied characteristics of prospective FMT donors that are predictive of passing a public stool bank's current screening process. FMT donor characteristics included in this analysis were donor's gender, age, body mass index, frequency of bowel movements, diet, alcohol and tobacco use, and seasonality based on the screening date. This study also characterized the 16S rRNA intestinal microbiome composition of stool from FMT donors that passed from the prospective to the active donor phase of the stool bank's FMT screening program. Hence, the second purpose of this study was to determine whether the microbial diversity of active FMT donor stool, as measured by 16S ribosomal ribonucleic acid, is related to FMT clinical efficacy, as measured by the recipient's clinical cure. The data for these analyses were on file and sourced from OpenBiome and were made available to the researcher.

Specific Aims

The specific aims of this secondary data analysis were to:

1. Identify additional, previously unstudied characteristics of prospective FMT donors that are predictive of passing a public stool bank's current screening process.
2. Determine whether the microbial diversity of active FMT donor stool, as measured by 16S ribosomal ribonucleic acid, is related to FMT clinical efficacy, as measured by the recipient's clinical cure.

Definition of Terms

The following section clarifies important terms that were relevant to the implementation of this study. Theoretical and operational definitions for prospective FMT donor characteristics used for this study are defined in the methods section.

Prospective Donor. An individual who has expressed interest in donating fecal material for use in FMT and has demonstrated a willingness to complete the FMT donor screening process (World Health Organization, 2012).

Deferred Donor. A prospective donor who was postponed from donating material due to not passing at least one component of the FMT donor screening program, during either the health assessment or laboratory screening, and was deemed ineligible to donate fecal material to treat recurrent CDI at the time of screening (World Health Organization, 2012).

Active Donor. A prospective donor who has passed all the criteria of the donor screening program, including both the health assessment and laboratory screening, and has been deemed eligible to donate fecal material to treat recurrent CDI at the time of screening. Active donors are assumed to have a low risk of transmitting diseases and have been cleared to contribute biological tissue for transfusion or transplantation (World Health Organization, 2012).

Microbiota. The collection of microorganisms that colonize a particular environment, such as the intestinal microbiota (Eloe-Fadrosh & Rasko, 2013).

Microbiome. The collective genetic make-up of the bacterial inhabitants or microbiota that constitute a particular habitat or ecosystem, such as the intestinal microbiome (Eloe-Fadrosh & Rasko, 2013).

Phylum. A top level taxonomic rank or relative level for bacteria in the biological classification system. Immediately follows the Bacteria/Eubacteria kingdom and is above bacterial class (Tyler, Smith, & Silverberg, 2014).

Species. The basic unit or taxonomic rank for bacteria in the biological classification system based on a microbe's 16S ribosomal ribonucleic acid (rRNA) sequence (Tyler et al., 2014).

16S Ribosomal Ribonucleic Acid (rRNA). The highly conserved molecular genetic sequence component of the 30S ribosome that allows bacteria to be to identified and categorized into their representative phylogenetic classifications (Tyler et al., 2014).

Alpha Diversity (a.k.a. Microbial Diversity, Bacterial Diversity, Diversity). The mean diversity of the microbiome within a sample as measured by the richness and evenness of the microbiota (Tyler et al., 2014).

Beta Diversity. The degree of bacterial differentiation within the intestinal microbiome as measured by ratios, principal coordinates, and analysis plots (Tyler et al., 2014).

Dysbiosis. Defined as an imbalance in the microbiota resulting in diminished colonization by commensal bacteria that are normally present and an overabundance in colonization by potentially pathogenic non-commensal bacteria (Eloe-Fadrosh & Rasko, 2013).

Clinical Efficacy. The clinical efficacy in terms of recurrent CDI is the overall measure in percent of how effective an active donor's FMT material is at preventing CDI recurrence within the eight weeks post-intervention. The rate is most frequently provided as an overall percentage for all FMT treatments or as an individual percentage by FMT donor (Osman et al., 2016).

Clinical Cure. Clinical cure in the case of recurrent CDI is defined as the complete resolution of recurrent CDI symptoms for at least eight weeks post-procedure. Clinical cure is defined as a yes or no dichotomous outcome variable and established during the 8-week post-procedure physician follow-up assessment (Cohen et al., 2010; Surawicz et al., 2013).

Assumptions Based on Existing Knowledge

The following assumptions underpin this study:

1. Prospective FMT donors accurately reported demographic and clinical data and accurately answered the screening questions.
2. Active FMT donors who passed the prospective FMT screening program donated the fecal material collected.
3. The 16S rRNA methodology used to characterize the intestinal microbiome of FMT donors was the best technological option for differentiating between bacterial species within the FMT donor fecal material.
4. The FMT donor data were accurately recorded in the stool bank data repositories by the study staff.
5. The current approach used by the public stool bank for screening FMT donors effectively selected donors with safe and efficacious fecal material.

Chapter 2. Conceptual Framework and Review of the Literature

The previous chapter illustrated the significance of *Clostridium difficile* infections (CDI) and described how this study aimed to build knowledge that would improve the efficiency of treating this serious patient and public health problem. The following chapter presents the theoretical and conceptual framework as proposed by Eloe-Fadrosh and Rasko (2013) that was used to guide this study. Attention is placed on explaining the relationships among the human host, environment, and microbiota as they are expressed in healthy patients and in patients suffering from recurrent *Clostridium difficile* infections. The rationale for choosing this conceptual framework will be illustrated through its application within the pathophysiological process observed before and after fecal microbiota treatment (FMT) delivery in patients suffering from recurrent CDI. This chapter also aims to provide an overview of the existing literature and evolution of FMT as they relate to the variables of interest in this study. The literature review will highlight areas where the research currently falls short both conceptually and methodologically and how this study aimed to address these gaps.

Conceptual Framework

The theoretical and conceptual underpinnings for this study were drawn from a conceptual framework presented by Eloe-Fadrosh and Rasko (2013), which explains the relationships among host, environmental, and microbiota features (Figure 1). This ecologically based framework provides a model for explaining how the characteristics of donors of FMT material, along the donor's environment and intestinal microbiome, interact and influence each other to promote a symbiotic and homeostatic intestinal fecal microbiome conducive to host health. This model also explains how colonization by *Clostridium difficile* bacteria perpetuates host disease in affected patients and presents a plausible framework for how fecal transplant

material from healthy donors is likely to be effective in preventing the recurrence of CDI in patient recipients.

Figure 1.
Conceptual underpinnings

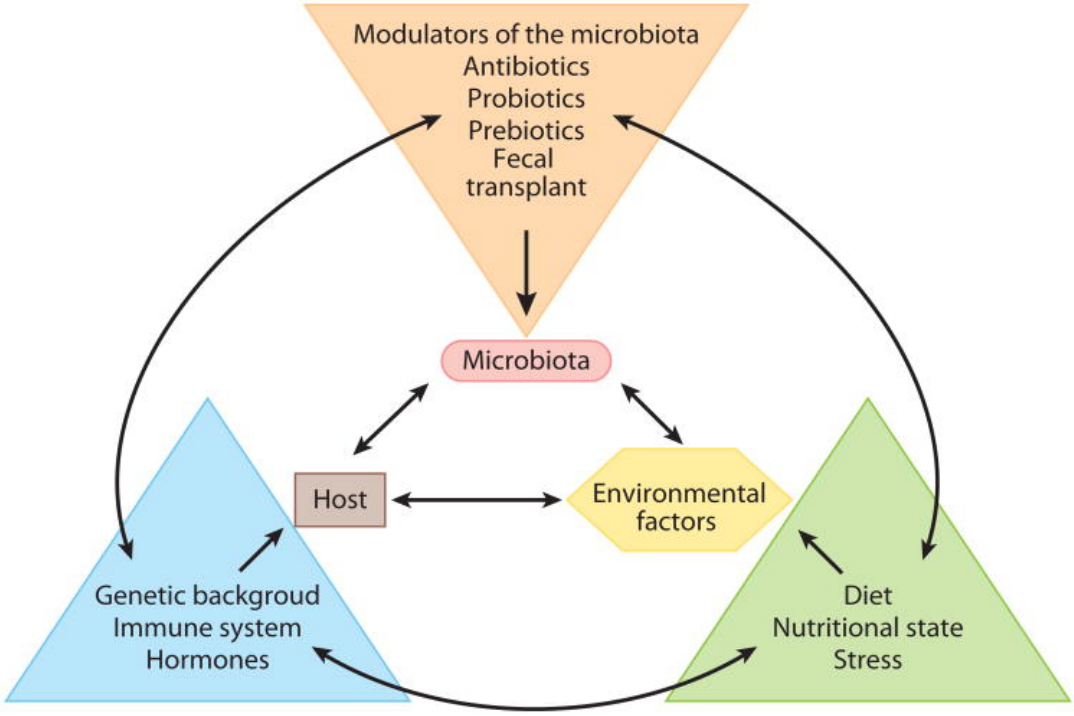


Figure 1. Framework demonstrating interactions among host, environmental, and microbiota features that promote patient health or lead to disease. Adapted from Eloe-Fadrosh, E. A., & Rasko, D. A. (2013). The human microbiome: From symbiosis to pathogenesis. Annual Review of Medicine, 64, 145-163. doi:10.1146/annurev-med-010312-133513.

Interactions among Host, Environmental, and Microbiota Features. Emerging research in ecology and medicine has demonstrated how the perspective of the host as the dominant influence in health and disease has evolved to view the host as one of several active participants. Equally important are environmental and microbiota features that collaboratively interact and work together with the host in a complex ecosystem to promote health or influence disease development (Costello, Stagaman, Dethlefsen, Bohannan, & Relman, 2012). Rather than view each feature independently, the ecological perspective perceives all features as having a

collective role in promoting host health or disease that is uniquely shaped by the factors that impact each of them directly.

The host's environment provides frequent exposure to factors that influence the composition of the intestinal microbiome on a daily basis. Host diet and nutrient consumption in particular, play an important role in defining microbial exposure and nutritional state. What the host consumes is further influenced by factors including the availability and accessibility of food and nutrients. In turn, these factors shape the intestinal ecosystem by providing microbial exposure and nutrient support that results in a tailored intestinal microbiome. Researchers are finding that colonization of the human host intestinal microbiome is not solely dependent on the environment and not entirely stochastic (Christian, Whitaker, & Clay, 2015; Costello et al., 2012). Alternatively, host specific factors, including genetics, gender, and stress influence the intestinal ecosystem and promote colonization by certain microbes to a much greater extent than previously perceived.

The host, environment, and microbiota features interact holistically to contribute to host health and disease prevention. One of the primary ways in which this complex ecosystem helps promote host health is by facilitating intestinal colonization resistance. Homeostatic community assemblage of the intestinal microbiome creates bacterial interference, which allows resident microbiota to hinder localized growth by potentially infectious exogenous microorganisms. The process of developing intestinal colonization resistance is highly influenced by the host, environment, and microbiota features.

Establishment of Intestinal Colonization Resistance. The intestinal microbiota is most recognized for helping to facilitate digestion and nutrition. However, research has shown that these intestinal microbes also play a critical role in health promotion and infectious disease

prevention through the development of colonization resistance (Tourneur & Chassin, 2013). Colonization resistance is established during community assemblage as intestinal bacteria interact with the human host to trigger immune mediated responses. This bi-directional communication leads to either host recognition of commensal bacteria or a host mediated immune response to pathogens. In turn, this process facilitates an intestinal ecosystem that promotes optimal colonization by microbes capable of independently mounting a localized competitive bacterial response and providing interference against invasive foreign organisms (Tourneur & Chassin, 2013).

Community assemblage of the intestinal microbiome is most pronounced during birth when newborns are exposed through the environment to a significant number of microbes. These microorganisms shape the newborn intestinal microbiome and are influenced by whether the birth was vaginal or by cesarean section (Dominguez-Bello et al., 2010). While the intestinal microbiome continues to fluctuate during infancy in response to distinct events, such as development related changes in diet, the intestinal microbiome generally stabilizes by two years of age to reflect an adult-like composition (Eloe-Fadrosh & Rasko, 2013). After that, colonization resistance is generally stable but continues to be influenced by aberrations in microbiome, host, and environmental interactions that directly impact host health.

The intestinal microbiome promotes colonization resistance throughout the host's life by utilizing several bacterial interference approaches. In addition to triggering the host's innate immune system in response to new or foreign pathogens (Corthesy, Gaskins, & Mercenier, 2007), bacteria that typically comprise the intestinal microbiome of healthy individuals compete directly with foreign microbes for scarce energy, resources, and space (Bernet-Camard et al., 1997). These processes skew the intestinal ecosystem toward an environment that impedes

colonization by opportunistic microbes (Chan, Reid, Irvin, Bruce, & Costerton, 1985). Commensal bacteria are also known to produce bactericidal proteins within the intestinal lumen that act locally to prevent colonization by competing bacteria (Brook, 1999). Once a healthy homeostatic intestinal ecosystem is in place, the symbiotic intestinal relationship between host and microbiome promotes a competitive and diverse microecosystem comprised of over 1,000 different species and dominated by bacteria from several different phyla including *Firmicutes*, *Bacteroides*, and *Actinobacteria* (Tuddenham & Sears, 2015). This diversity makes it challenging for any single potentially antagonistic organism to predominate and cause symptomatic intestinal disease in healthy hosts (Falagas, Rafailidis, & Makris, 2008).

In general, intestinal microbiome colonization and immune system development and function are robust to environmental insult. Symbiotic interactions among host, environment, and microbial features are designed to support intestinal colonization by a diverse microbiome, which in turn leads to colonization resistance and bacterial interference against potentially infectious pathogens. However, when disruptions to the symbiotic intestinal ecosystem occur, intestinal homeostasis and immune expression are hindered and can result in intestinal dysbiosis and disease development. Factors such as phenotypic expression of heritable diseases, extreme changes in diet, or recent use of antibiotics can lead to perturbations in the intestinal microbiome. These factors can also trigger a cascade of disruptions that span host, environment, and microbial features resulting in intestinal dysbiosis and increased susceptibility to disease. For example, a tooth abscess in the host may impact nutritional intake and require antibiotic intervention leading to disruption in intestinal colonization resistance. Sufficient perturbations to the intestinal microbiome, such as those observed with antibiotic use, may result in suboptimal bacterial

interference. This in turn places the host at an increased risk for colonization by opportunistic pathogens and disease development, such as *Clostridium difficile* related infections.

Application of Conceptual Framework to *Clostridium difficile* Infections

The *Clostridium difficile* bacterium is a pernicious pathogen that leads to infectious diarrhea and colitis. Although infections caused by the *Clostridium difficile* bacterium are not new to healthcare, the incidence and severity of CDI have increased over the past decade (Borgia et al., 2015; Khoruts & Weingarden, 2014). Each year, over 453,000 new cases of CDI are diagnosed in the United States (U.S.), of which approximately 83,000 are repeat occurrences due to failed antibiotic treatment (Lessa et al., 2015). Primary risk factors for CDI include physiological states or medical interventions that decrease intestinal bacterial diversity, such as antimicrobial use (Khanna et al., 2012). These risk factors impact host immune response, reduce bacterial interference, and present an opportunity for colonization by potentially pathogenic *Clostridium difficile* spores.

At any point in time, approximately 1-3% of the adult population is colonized by *Clostridium difficile* microbes (Ozaki et al., 2004). However, as an opportunistic bacterium, *Clostridium difficile* bacteria do not typically cause disease in healthy adults with intact colonization resistance (Rea et al., 2012). Bacterial interference mounted by resident microbes in a healthy and diverse intestinal ecosystem hampers active colonization by *Clostridium difficile* bacteria and its ability to multiply and cause a disease state. In spore form however, *Clostridium difficile* is resistant to intestinal bacterial interference, antibiotics, and most hospital disinfectants. This allows *Clostridium difficile* spores to reside dormant in the intestines and on surfaces for prolonged periods of time, in wait for an environment that is conducive to growth (Khan &

Cheesbrough, 2003). Figure 2 provides an overview of the cycle of CDI infection and recurrence within the context of intestinal microbiome colonization resistance.

Figure 2.
Cycle of *Clostridium difficile* infection and recurrence

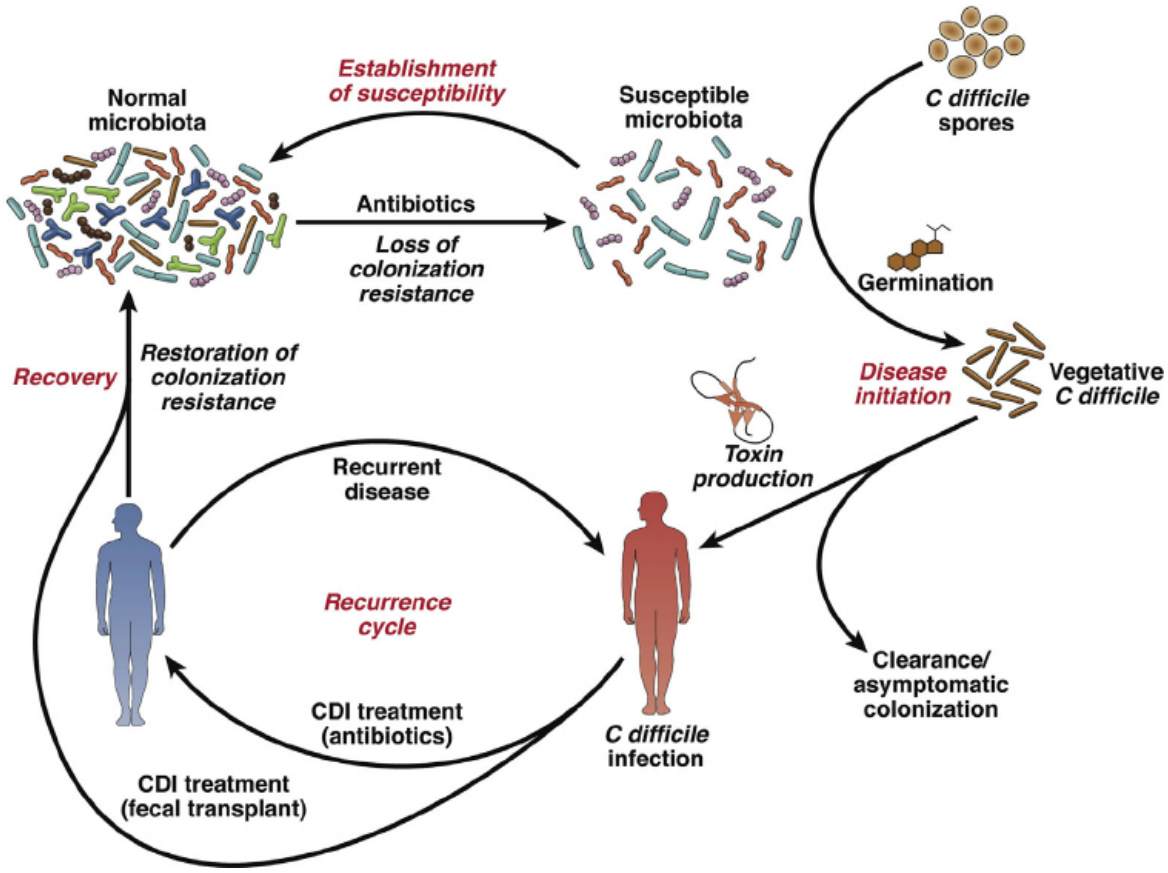


Figure 2. Repeated antibiotic use alters the patient's intestinal microbiome, setting up an environment that is permissive to colonization and repeat infection by *Clostridium difficile*. Adapted from Britton and Young (2014). Role of the intestinal microbiota in resistance to colonization by *Clostridium difficile*. Gastroenterology, 146(6), 1547-1553. doi:10.1053/j.gastro.2014.01.059

Disruption of the normal intestinal microbiome as a result of medical interventions, such as antibiotics, creates dysbiosis of the intestinal microbiota. This breakdown in colonization resistance places the host at increased risk for colonization by opportunistic pathogens. *Clostridium difficile*, a bacillus shown to survive in spore form for up to three months on surfaces, is transmitted via the fecal-oral route and is capable of surviving the acidity of the

stomach. Infectious diarrhea and colitis occur when *Clostridium difficile* spores that have been ingested by an individual encounter a perturbed intestinal microbiome where bacterial interference is compromised. Under this opportune condition, the *Clostridium difficile* spores germinate and quickly colonize the large intestine with little resistance. *Clostridium difficile* begins producing endotoxins that further skew the intestinal environment toward its benefit, creating an environment unfavorable to colonization by commensal bacteria. The resulting infection can range from mild diarrhea and abdominal pain to severe pseudomembranous colitis, toxic megacolon, sepsis, and death.

For 94% of inpatient and 78% of outpatient CDI cases, exposure to antibiotics occurs within the 90 days prior to diagnosis (Weissman & Coyle, 2012). Despite the association, antibiotics have remained one of the few available options for treating infections caused by *Clostridium difficile* bacteria. The aim of antibiotic use is to suppress the growth of *Clostridium difficile* bacteria and reduce the production of endotoxins that produce localized inflammation and exacerbate CDI. For approximately 80% of patients with a first episode of CDI, antibiotics are successful at eradicating the disease (Kelly, de Leon, & Jasutkar, 2012). However, antibiotics also indiscriminately target critical commensal intestinal bacteria, making it difficult for the host to restore colonization resistance. Because antibiotics are not capable of clearing residual *Clostridium difficile* spores, these spores remain dormant in the intestines and convert into the disease-causing vegetative state once the antibiotic has been cleared from the patient's system. This allows the *Clostridium difficile* bacterium to take advantage of the sustained dysbiotic intestinal environment and cause recurrence before commensal bacteria can repopulate and restore colonization resistance (McFarland, 2012). For the remaining 20% of first episode CDI patients treated with antibiotics, recurrence of disease occurs within eight weeks after antibiotic

treatment and is associated with a notably higher 40% chance of recurrence driven by prolonged intestinal dysbiosis (Kelly et al., 2012). The risk for recurrence rises further to over 60% for patients who fail two or more rounds of CDI antibiotic based interventions.

Use of repetitive and continuous antibiotic regimens to treat recurrent CDI leads to a notable loss of intestinal microbial diversity and colonization resistance. After repeated use, patients treated with antibiotics for recurrent CDI develop an intestinal microbiome that is perpetually reduced in overall intestinal microbial richness and diversity. Most notably, commensal bacteria from the *Bacteroides* and *Firmicutes* phyla that are usually plentiful in a healthy intestinal ecosystem are significantly reduced as a group in patients with recurrent CDI (Chang et al., 2008). As a result, repetitive use of antibiotics to treat recurrent CDI irreparably perturbs overall microbial density and alters the entire intestinal ecosystem long-term (Vincent & Manges, 2015), hindering the ability of the host to restore colonization resistance and perpetuating the cycle of recurrence observed with CDI.

Restoration of Intestinal Colonization Resistance by Fecal Transplants. Although fecal material has been used for centuries to treat intestinal related diseases (Zhang, Luo, Shi, Fan, & Ji, 2012), the use of human fecal material to successfully cure pseudomembranous colitis associated with CDI was first reported in the medical literature in 1958 (Eiseman, Silen, Bascom, & Kauvar, 1958). More recently, with the advent of genetics sequencing technologies and increased knowledge about the human intestinal microbiome, FMT has emerged as a viable restorative treatment for patients who experience recurrent CDI. The primary aim of FMT is to impede *Clostridium difficile* reinfection by quickly restoring intestinal colonization resistance post-antibiotic treatment. Fecal transplants encompass the infusion of homogenized and unfiltered human fecal material from either directed donors i.e., a donor known to the patient, or

anonymous donors. The fecal transplant preparation is infused directly into the colon or delivered to the small intestines through upper endoscopy or capsules within two to three days after completion of antibiotics.

Post-transplant, the recipient's intestinal microbiome quickly repopulates and begins to resemble a diverse community composed of both the recipient and donor's microbiome (Seekatz et al., 2014). By 14 days post-transplant, bacteria from the *Bacteroides* phylum and butyrate producing bacteria known to be associated with intestinal health are found to dominate the intestines (Khoruts, Dicksved, Jansson, & Sadowsky, 2010). Corresponding with intestinal re-colonization, recipients report gradual improvement and subsequent resolution in diarrhea, bloating, and abdominal pain within one to four weeks of FMT treatment (Brandt et al., 2012). Recipients that experienced fatigue, anorexia, and loss of appetite during active CDI report improvement in energy, the restoration of appetite, and resumption of a normal diet, all of which further support overall well-being and host health. The competitive commensal microbes introduced by fecal material from donors restore colonization resistance and inhibits the growth of *Clostridium difficile* bacteria and prevent reinfection. In association with the restoration of overall host well-being and environmental features that promote host health, commensal microbes restore long-term intestinal microbial stability that results in symptom resolution and clinical cure.

Review of the Literature

Clinical Efficacy of Fecal Microbiota Transplant. Since 1958, several case series, meta-analyses, and two randomized controlled clinical trials have published evidence supporting the clinical efficacy of FMT in preventing recurrent CDI (Brandt et al., 2012; Kassam et al., 2013). More recently in 2013, van Nood and colleagues presented results from one of two

randomized controlled trials comparing the clinical efficacy of fresh fecal microbiota preparations from anonymous donors with the current standard of care antibiotic treatment for recurrent CDI. Interim data analysis suggested resolution rates of up to 94% in the FMT intervention group, compared to 31% or less in the standard of care groups, prompting early termination of the clinical trial (van Nood, Vrieze, et al., 2013). In 2014, Youngster and colleagues utilized frozen fecal material from anonymous donors to successfully treat patients with recurrent CDI demonstrating a clinical efficacy rate of 90% (Youngster et al., 2014). Smaller randomized controlled studies performed in notoriously difficult to treat population subgroups, such as immunocompromised or severe-complicated recurrent CDI, similarly reported high levels of clinical efficacy (Kelly et al., 2014).

While the rates of clinical efficacy for FMT reported in the literature can be high, FMT has been found to fail to prevent the recurrence of CDI in anywhere from 2.1% to 35.7% of patient cases (Fischer et al., 2016; Kassam et al., 2013). In a smaller percentage of patients, FMT is entirely ineffectual in preventing CDI despite repeated treatments (Fischer et al., 2016). The explanation for the less than 100% clinical efficacy rate of FMT in preventing CDI recurrence is uncertain and published literature that purports to explain this finding is limited. However, clinicians and researchers postulate that the rationale for the failure may be multifactorial, whereby donor features such as intestinal microbiome composition may be influential. In a retrospective cohort study of 201 subjects conducted in the U.S. in which 12.4% patients experienced CDI recurrence post procedure, FMT failure was found to be positively correlated with the number of FMTs received. This finding indicated that as the number of FMT treatments increased, the higher the likelihood of CDI recurrence (Meighani et al., 2016). This finding prompted the investigators to advise that a second FMT using an alternative donor be considered

in the case of FMT failure. Other smaller studies have similarly found that a second FMT treatment utilizing material from a donor presumably possessing a different intestinal microbial composition has resulted in CDI cure in patients where recurrence of CDI occurred post-procedure (van Nood, Dijkgraaf, & Keller, 2013). Findings presented by Fischer and colleagues (2016) identified a number of patient recipient factors that appeared to be predictive of FMT failure, including the patient's severity of CDI and FMT performed as an inpatient. However, no donor characteristics were considered as part of this study, which was noted by the investigators to be a limitation. Although the findings of these studies were informative to clinical practice, neither FMT donor characteristics nor the intestinal microbiome composition of the donor were evaluated to assess their potential impact on the clinical efficacy of FMT.

Despite the suggestions in the literature that donor intestinal microbiome composition and diversity may increase the clinical efficacy of FMT, little research has been done to address this gap in knowledge. The difficulty in fully understanding potential contributors to FMT failure was echoed in Kassam and colleagues' meta-analysis on the use of FMT in the treatment of recurrent CDI (Kassam et al., 2013). While the review found no clinical difference in overall clinical efficacy rates between anonymous versus directed donors, the authors did find significant disparity in the donor screening approach used by clinicians and noted that this may account for the lower clinical efficacy rates observed in some studies. As part of the conclusion, the investigators reiterated the need for guidance and greater standardization of donor screening to ensure the maximum likelihood of procuring both safe and efficacious donor fecal material.

Introduction of Fecal Microbiota Donor Screening Programs. Because of the high rates of FMT clinical efficacy and clinical expert testimony demonstrating patient need for FMT, in 2013 the U.S. Food and Drug Administration (FDA) chose to take an enforcement discretion

approach to regulating the use of FMT despite the U.S. FDA's classification of FMT as an experimental drug. The policy effectively allows clinicians to administer FMT for recurrent CDI as long as the patient is properly consented. As the use of FMT has grown under this permissive regulatory environment, so has the importance of finding suitable, healthy FMT donors that can contribute clinically efficacious fecal material. However, the process to implement such a program by healthcare facilities, public stool banks, and clinicians has been difficult and unpredictable.

The inconsistencies encountered among FMT donor screening programs create a significant impediment to FMT for both clinicians and patients. Nurses have an important role in the execution of many of the FMT donor screening program operational procedures and FMT delivery. These activities include championing institutional adoption, facilitating donor assessments and laboratory screenings, reviewing donor eligibility, procuring fecal material, and assisting in endoscopy or providing patient treatment directly through enema or capsules (Samuel et al., 2014). A recent study by Brumbaugh and colleagues (2017) presented the findings of a successful implementation of a nurse-led intragastric pediatric FMT program, demonstrating both economic benefits and efficacious FMT outcomes in line with the literature.

Because FMT is a relatively new procedure, however, most healthcare facilities lack the protocols needed to handle the donor screening and procurement of fecal material. The challenges with implementing FMT as a treatment option were elucidated by a nurse who chronicled the steps she went through to facilitate the first FMT within her hospital (Myers, 2015). Among several barriers identified, a knowledge deficit about effective donor screening criteria and burdensome U.S. FDA recommendations were highlighted. As a result, the methods utilized by healthcare facilities, researchers, and stool banks to develop FMT donor screening

programs have varied widely from simple clinician-guided individual donor screenings to full ethics board approved investigational drug research protocols (Bafeta, Yavchitz, Riveros, Batista, & Ravaud, 2017). Referred to as "heterogeneity in FMT protocols," Brumbaugh and colleagues acknowledged this phenomenon as a significant gap in clinical care and opportunity for improvement (Brumbaugh et al., 2017).

Adoption of a Risk Reduction Based Approach to Donor Screening. Much of the uncertainty and disparity across FMT programs has been driven by a lack of informative research on donor features, such as donor age or the donor microbiome, that are most pertinent to assess during the screening process. A recently published systematic review by Bafeta and colleagues (2017) on FMT donor screening approaches utilized by FMT clinical research trials substantiated the persistent lack of description of methodological components in sourcing fecal material (up to 89% of studies reviewed) and general guidance on donor screening methods (up to 98%). To address the gap in guidance, in 2011 the American Gastroenterology Association (AGA) FMT Working Group, in collaboration with the U.S. FDA, published a document outlining suggested minimum criteria to utilize in the screening of donors and administration of FMT within clinical practice (Bakken et al., 2011). The group suggested the adoption of a risk reduction approach to donor screening, similar to the approached presently utilized by blood banks (Center for Biologics Evaluation and Research, 2007) and in compliance with the U.S. FDA Title 21 Code of Federal Regulations (CFR). Under Title 21 CFR, subpart 1271, clinicians that screen biological material donors, such as blood and stool, must ensure that only material that is free of infectious and communicable disease is used. As such, donors are required to be healthy, at low risk for infectious disease, and demonstrate no clinical disease symptoms (Department of Health and Human Services, 2017). Noting the lack of data to support the recommendation for

considering other donor factors in the screening process, such as donor fecal microbiome composition and diversity, the American Association of Blood Bank's (AABB) Donor Health Questionnaire (DHQ) and laboratory based testing protocols were proposed as a template to evaluate prospective FMT donors (Bakken et al., 2011).

The primary rationale for utilizing a risk reduction approach to screening prospective FMT donors is to provide a way for nurses and other clinicians to identify as early as possible any risks for transmitting an infectious disease through donor stool. To accomplish this goal, clinicians are advised to select donors that are in good health and demonstrate a near zero risk for infectious disease. As such, every step in the risk reduction based prospective FMT donor screening program is designed to test donor characteristics against an exclusionary criterion. For example, like the blood donor risk reduction model, specific social and lifestyle behaviors that may indicate a donor has a higher risk for infectious or communicable disease, such as recent travel to countries where communicable diseases are prevalent, use of illicit drugs, and high-risk sexual behaviors, are also considered exclusionary under the FMT donor screening program.

Prospective FMT donors however, undergo more extensive in-person screenings than blood donors and meet with multiple clinicians, including nurses, who assist in making the determination as to the suitability of the donor's fecal material. For example, FMT donors with gastrointestinal issues (i.e., recurrent diarrhea or constipation) or recent clinical interventions that are known to perturb the intestinal microbiome (i.e., recent use of antibiotics) are excluded. Although long-term correlational data about the role of the intestinal microbiome on health and disease is limited, there is evidence to suggest that a dysbiotic intestinal microbiome is associated with certain chronic diseases, such as inflammatory bowel disease (IBD), intestinal malignancy, and obesity (Penders, Stobberingh, van den Brandt, & Thijs, 2007; D. A. Peterson,

Frank, Pace, & Gordon, 2008; Turnbaugh & Gordon, 2009). As such, prospective FMT donors with any history or clinical evidence for chronic gastrointestinal diseases are excluded based on the assumption that the dysbiotic fecal material may trigger the development of these chronic diseases in the recipient of the transplant material (Turnbaugh, Backhed, Fulton, & Gordon, 2008). Upon completion of the stool bank DHQ, clinicians perform in-person health exams and clinical interviews to ensure the accuracy of the donor's answers to the stool bank DHQ. Once a prospective FMT donor passes the risk assessment portion of the screening, prospective FMT donors undergo serological and stool laboratory testing (described in more detail in Chapter 3) for infectious diseases that may be transmissible through stool.

Fecal Donor Screening Pass and Deferral Rates. While the proposed risk reduction model, Blood Bank DHQ, and blood bank laboratory testing protocols provided a starting point for screening prospective FMT donors, the FMT screening program quickly became overly burdensome and inefficient for nurses and other clinicians to utilize due to the inherent risks associated with transplanting unprocessed fecal material. The use of this highly conservative, expansive, and logistically challenging risk reduction approach for assessing the suitability of prospective FMT donors is time-consuming, conducted over multiple in-person visits, requires completion of extensive laboratory testing, and takes several weeks to complete (Bakken et al., 2011). Ultimately, the increased complexity of the FMT donor screening programs when applied to the general population reports donor pass rates between 4-10%, and deferral rates between 90-96% (L. J. Burns et al., 2015; Paramsothy et al., 2015). While the risk reduction approach was important to provide a process for screening out infectious diseases for which suitable tests are not available, the result is a significant reduction in the qualified FMT donor pool. Despite this

situation, no research has been conducted to identify the predictors of FMT donor passage and failures to improve the process for finding healthy stool donors.

Using Blood Banks as a Model to Improve FMT Donor Pass Rates. Since adoption of the risk reduction based approach, blood banks continue to utilize and improve the model to ensure the safety of blood transfusion material (World Health Organization, 2012). However, the deferral and pass rates for blood donors is notably lower at 12.8%, with an 87.2% pass rate, compared to fecal donor pass rates (Zou et al., 2008). To keep this deferral rate low, blood banks regularly monitor and analyze prospective donor characteristics that may be predictive of passing or deferral. This information is subsequently used to improve the recruitment, tailor the blood donor screening processes, and increase donor pass rates when implemented.

In addition to data on deferral rationales, donation frequency, and donor attrition, blood banks regularly monitor various prospective donor factors to inform and target donor recruitment. These include factors such as gender, age, blood type, and time of donation, as well as certain health and lifestyle factors, such as dietary habits, body mass index, and drug and alcohol use (World Health Organization, 2012). Data collected about prospective blood donor characteristics, behaviors, and outcomes are analyzed and used to support continuous improvement in operations, monitor adherence to standard operating procedures, and ensure the clinical efficacy of blood transfusions. For example, an evaluation of European blood donors by Bani and Giussani (2010) demonstrated that gender had a significant impact on prospective donor motivation, deferral, and donor related adverse events. These findings prompted tailoring of the prospective blood donor recruitment and screening programs to properly evaluate prospective female donors for specific risk factors. Female donors found to have a higher risk of deferral from whole blood donations were directed to alternative types of donations that were

more suitable with their hematological parameters. In turn, these operational changes reduced deferral rates and costs associated with screening prospective blood donors. Ultimately, donor data allow blood banks to efficiently recruit donors who are more likely to pass the screening and minimize the cost and delays resulting from donor deferrals.

Challenges with Improving Fecal Donor Screening Programs. Despite its limitations in screening FMT donors, the blood banks' risk reduction model has remained the principal method for screening prospective FMT donors (Kapel et al., 2014; Kelly et al., 2012). Further, very little research has been conducted to date to improve the FMT donor screening process. The primary challenge of studying prospective FMT donor characteristics, 16S ribosomal ribonucleic acid (rRNA) intestinal microbiome composition, and related clinical efficacy rates to improve the procurement of suitable fecal material has been the limited pool of information about qualifying donors. Because the directed donor approach requires only one donor and one universal FMT donor can treat many patients, sufficient aggregation of data to support meaningful inquiry has been limited. Further, despite reductions in the cost of genetic sequencing technologies, microbiome sequencing adds to the already high out-of-pocket procedural costs associated with screening FMT donors. Because the use of microbiome sequencing is not currently supported by the literature, physicians and stool banks are reluctant to incorporate it into donor screening despite its potential to improve the clinical efficacy of FMT in patient recipients. As a result, these barriers have hindered progress and process improvements to the FMT donor screening programs currently utilized by clinicians and stool banks.

Variables of Interest to this Study

FMT donor characteristics that impact the host, environment, and microbiota features presented within the ecological based conceptual framework by Eloe-Fadrosh and colleagues (2013) were evaluated as part of this study to assess whether this information can be utilized to improve the FMT donor screening process to better identify healthy donors. Specific FMT donor characteristics of concern to this study were: 1) donor host factors including gender, age, body mass index (BMI), and frequency of bowel movements; 2) environmental factors including diet, alcohol and tobacco use and seasonality based on screen date; and 3) 16S rRNA based intestinal microbiome diversity. The aforementioned variables have been supported by the literature to influence an individual's health and microbiome composition and are thus potentially predictive of donor pass rates and FMT clinical efficacy. The following section outlines the current state of the science as it pertains to each of the FMT donor characteristics of interest to this study.

Host Features.

Gender. Although research supporting an interaction between gender-related hormones and intestinal microbial composition is still emerging, there is strong evidence to suggest that gender-related differences in health and the intestinal microbiome exist. The difference is thought to be due in part to the bi-directional modulation of sex hormones by the host and intestinal microbiome (Garcia-Gomez, Gonzalez-Pedrajo, & Camacho-Arroyo, 2013). The evolving maternal intestinal microbiome during pregnancy demonstrates how rising estrogen levels during the third trimester lead to a distinct and less diverse intestinal microbiome independent of health status and dominated by *Proteobacteria* and *Actinobacteria* (Koren et al., 2012). More recently, a study looking at gender differences in intestinal microbiome within the context of BMI demonstrated that the intestinal microbiome between matched subjects differed

by gender at the bacterial phyla, genus, and species levels (Haro et al., 2016). Despite the lack of a clear causal explanation between gender and intestinal microbiome mediated disease, the research provides evidence to suggest that gender driven differences in the intestinal microbiome exist and that these differences uniquely shape 16S rRNA microbiome composition and host health.

No research to date has been conducted to assess whether one particular gender demonstrates a higher frequency in passing the FMT donor screening, possesses a distinct 16S rRNA intestinal microbiome, or offers better clinical efficacy rates in treating recurrent CDI. Alternatively, female donors who are either pregnant or breastfeeding are generally excluded from donating due to the observed changes in 16S rRNA intestinal microbiome composition associated with fluctuations in hormone levels. The female gender has also been associated with a higher risk for certain functional gastrointestinal disorders including irritable bowel syndrome (IBS), recurrent diarrhea and constipation (Bakken et al., 2011; Houghton et al., 2016), and celiac disease (Lo, Sano, Lebwohl, Diamond, & Green, 2003), suggesting that female FMT donors may be less likely to pass the FMT donor screening. Because individuals with symptoms consistent with gastrointestinal disorders may also present with a perturbed intestinal microbiome in the absence of a formal diagnosis, these FMT donors are consequentially excluded from donating fecal material. In general, the risk reduction approach excludes certain health related factors that appear to affect women disproportionately. As such, there is the possibility that female FMT donors may be less represented in the FMT donor pool and that the 16S rRNA intestinal microbiome composition for female FMT donors that do pass may be significantly different from that of male FMT donors. As such, this study postulated that identifying any gender related influences could provide valuable and pertinent information that

may improve FMT donor recruitment logistics, decrease screening costs, and promote consistency in donor microbiome diversity and FMT clinical efficacy rates.

Age. Studies on the intestinal microbiome in elderly individuals have demonstrated that as people age, there is a shift in colonization toward a less diverse microbiome composition (O'Toole & Jeffery, 2015). More specifically, elderly individuals tend to have higher relative proportions of *Bacteroidetes* (Hopkins, Sharp, & Macfarlane, 2002) when compared to young adults who are predominantly colonized with higher proportions of *Firmicutes* (Cresci & Bawden, 2015; Mariat et al., 2009; Nagao-Kitamoto, Kitamoto, Kuffa, & Kamada, 2016). While an exact mechanism or explanation for this shift in intestinal microbiome composition is not fully understood, the gradual evolution in microbiome composition is associated with an increasing risk for immunosenescence and declining cognitive function as one ages (Biagi et al., 2013; Guigoz, Dore, & Schiffrin, 2008; O'Sullivan et al., 2013). This shift in diversity also impacts colonization resistance over time, resulting in suboptimal bacterial interference in older individuals and thus increasing one's risk for opportunistic infections. In particular, *Clostridium difficile* related infections have been shown to disproportionately affect the elderly population in spite of one's health status (Loo et al., 2011).

While intestinal microbiome variability among individual adults has been found to differ, there is still uncertainty as to whether aging plays a role in 16S rRNA intestinal microbiome composition within the healthy adult population. This has led to the assumption that the intestinal microbiome composition and colonization of young to middle aged adults do not differ. Because intestinal microbiome related research has predominantly focused on the extreme ends of the age spectrum, namely the very young and elderly, there are few studies looking at within group age-related differences in the intestinal microbiome of healthy adults. However, a recent cohort study

of Japanese community dwelling individuals aimed to illustrate the evolution of the intestinal microbiome across different decades of life from infancy to elderly. That study demonstrated that the microbiome composition of the adult population as a group remained fairly stable (Odamaki et al., 2016). However, intestinal microbial diversity gradually decreased upon adulthood after eighteen years of age and with each decade of life.

FMT donors of advanced age are generally excluded from donating based on the increased risk for dysbiosis and disease found to be associated with the elderly. Due to differing interpretations of what constitutes advanced age within the context of health and intestinal microbiome related disease, FMT donors as young as 50 years of age may be excluded from donating fecal material. Further, the various demands and constraints inherent in the FMT donor risk reduction approach and logistics of donating fecal material such as in-person health assessments, multiple laboratory-based screenings, and on-site or time-sensitive donations, may favor enrollment by a younger subset of the healthy adult population. Because the compositional stability of young to middle aged adults and the timing at which intestinal microbiome diversity evolves are not entirely understood within the context of host physiology and health, this study postulated that evaluating whether age is predictive of passing the FMT donor screening process could prove to be informative.

Body Mass Index. Healthcare providers routinely measure a person's BMI as a proxy for total body adiposity and disease risk. A healthy BMI is thought to range from 18kg/m^2 to less than 25kg/m^2, while a BMI between 25kg/m^2 and less than 30kg/m^2 is considered overweight. Generally, individuals with a BMI of 30kg/m^2 or greater are considered obese and have significantly elevated odds of prematurely developing chronic diseases, such as metabolic and cardiovascular related disease (Mokdad et al., 2003; Nuttall, 2015). Risk factors for a high BMI

include a sedentary lifestyle, poor dietary habits and nutrition, genetics, and certain medical conditions (Centers for Disease Control and Prevention, 2017a). Of concern in developed countries is the increase in sedentary lifestyles and poor dietary habits that are driving a growth in obesity rates. When caloric consumption is not offset with an increase in caloric expenditure, the host's increased access to nutrients allows cells to process and store greater amounts of energy as fat cells, resulting in excessive weight gain and increasing one's risk for obesity. While the health issues in obese individuals are more apparent, individuals who fall within the overweight category (BMI between 25 and 30kg/m^2) are also generally perceived to be at a heightened risk for long-term development of obesity related chronic diseases.

Recent research has proposed that diets high in fat and calories may lead to a dysbiotic intestinal ecosystem colonized by microbes that are conducive to weight gain (Ley et al., 2005). Using a humanized mouse model, Turnbaugh and colleagues (2008) were able to transplant fecal material from obese humans and replicate the diet-induced obesity effects of a high caloric, high fat diet on the mouse model. Upon introduction of the high caloric, high fat diet, mice transplanted with dysbiotic fecal material gained weight and continued to do so despite caloric reductions. Subsequent research on obesity conducted in human twins found reduced intestinal microbial diversity and phylum level dysbiosis in obese twins when compared with their lean counterparts (Turnbaugh et al., 2009). This difference was found to be significant and independent of innate host or environmental factors. The research suggests that intestinal colonization by a distinct set of microbes inherent in obese individuals may be the impetus for further sensitization of host metabolic pathways (Cox & Blaser, 2013) and result in further weight gain over time.

High caloric, high fat diets promote a shift in the intestinal ecosystem causing it to perpetuate the intestinal microbial dysbiosis that results in weight gain. Referred to as efficient harvesters, the intestinal microbes associated with obesity may serve as biomarkers and precursors for obesity-related chronic diseases. Because individuals with a BMI of 30kg/m^2 or greater are generally understood to have a heightened risk for intestinal dysbiosis and exhibit a host of disease symptoms, including diabetes, hypertension, and high cholesterol, FMT donors with a BMI of 30kg/m^2 or greater are excluded from donating fecal material. However, the risks for disease and intestinal dysbiosis in overweight individuals with BMI between 25kg/m^2 and less than 30kg/m^2 are not as well understood and less apparent. Because the BMI metric is an imperfect measure of host health, some individuals with a BMI between 25kg/m^2 and 30kg/m^2 may not accurately be classified as overweight or at risk for intestinal dysbiosis. Evaluating FMT donor data to assess for a BMI-related impact on deferral status would be valuable to help identify whether certain individuals are more likely to pass based on their BMI. Alternatively, overweight individuals may already possess an asymptomatic dysbiotic intestinal microbiome or present with mild obesity related symptoms that don't yet meet strict exclusion criteria. As such, this study postulated that evaluating the donor BMI with passing the prospective donor screening program, may provide a metric that would be more efficient at identifying healthy FMT donors earlier in the screening process.

Frequency of Bowel Movements. Transit time of fecal material through the intestines and fecal consistency have been shown to directly impact the type of microbiota that colonize the intestinal tract through access to nutrients and bacterial clearance (Tigchelaar et al., 2016; Vandeputte et al., 2016). Although abnormally pronounced during infectious diseases, healthy stool frequency, which serves as a proxy for intestinal transit time, can vary from as few as three

stools per week to three stools per day. In the absence of disease, frequency of bowel movements affects the length of time that fecal material remains in the intestines where nutrient breakdown and water absorption by the host and microbial inhabitants occur. Accelerated transit of fecal material can produce nutritional scarcity and disproportionately impair colonization by certain microbes. In a recent study, rapid fecal transit was found to selectively skew intestinal colonization toward fast growing bacteria, such as *Ruminococcaceae* and *Bacteroides* (Vandeputte et al., 2016). These bacteria were found to be capable of utilizing available nutrients more effectively allowing for greater colonization. As a result, intestinal microbial composition can differ markedly depending on the length of time that intestinal microbes have access to nutrients.

Depending on the fecal transit time, differentiated ecosystems may emerge that support colonization by distinct bacteria and influence microbiome composition and colonization resistance. Accelerated fecal transit increases bacterial clearance and washout. This hinders the ability of certain bacteria to adhere to and colonize the intestines. As a result, bacteria with a higher degree of intestinal wall adherence capabilities, such as that seen with *Prevotella*, are more likely to avoid bacterial washout (Vandeputte et al., 2016) and disproportionately colonize the intestines. Of concern is the finding that the intestinal microbiome of individuals with *Clostridium difficile* related infection and non-infectious diarrhea in healthy adults was found to be remarkably similar and indistinguishable in diversity (Schubert et al., 2014). These microbiome findings were in contrast to non-diarrheal healthy controls that maintained significantly different intestinal microbiomes. In particular, the *Clostridium difficile* bacterium is known to adhere well to intestinal epithelial cells. As such, the intestinal microbiome of individuals with more frequent bowel movements may be comprised of a dysbiotic intestinal

microbiome composition and reduced colonization resistance that places them at greater risk for infection by *Clostridium difficile*.

Despite the findings in the literature, transit time and stool consistency are not taken into consideration as potential influencers of the intestinal microbiome or host health during the FMT prospective donor screening process. Stool frequency and consistency, as measured by the Bristol Stool Scale, are collected as part of the FMT donor screening process. However, the aim of screening stool characteristics is to reduce the risk for a perturbed microbiome due to an infectious cause and uncover donors who have a history of chronic diarrhea or constipation related to an undiagnosed gastrointestinal disorder. Due to the lack of research in FMT donors, neither transit time nor stool consistency are currently utilized as exclusion criteria or potential predictors of target FMT donors with a diverse intestinal microbiome and potentially higher FMT clinical efficacy. As such, this study postulated that investigating the donor frequency of bowel movements could prove to be informative.

Environmental Features.

Diet. One of the beneficial functions of intestinal bacteria is the ability of these microbes to breakdown undigested foods and promote access and absorption to important nutrients (Turnbaugh & Gordon, 2009). The type of diet one consumes directly shapes the intestinal ecosystem and bacteria that inhabit the intestines long-term by promoting colonization by bacteria capable of digesting the nutrients that are consumed as part of the diet. This influence can be observed in the intestinal microbiome among individuals who consume animal-based versus predominantly carbohydrate- or plant-based diets. Diets of individuals who consume plant-based foods and exclude all animal products, referred to as vegetarians, tend to be higher in carbohydrates and fiber and lead to lower stool pH and lower counts of *Bifidobacterium*,

Bacteroides, *Escherichia coli* and *Enterobacteriaceae* (Zimmer et al., 2012). *Escherichia coli* and *Enterobacteriaceae* species are known to survive better in protein-rich, high pH intestinal ecosystems. As such, these bacteria tend to be more predominant in individuals that consume animal proteins. Whereas, individuals who consume primarily carbohydrate- or plant-based diets are predominantly colonized by the *Prevotella* enterotype (De Filippo et al., 2010; Wu et al., 2011).

Similar trends have been observed in individuals who follow non-traditional diets, which have gained in popularity over the past decade regardless of the rationale or health need, such as gluten-free diets. Research on individuals who adopted gluten-free diets for a four-week period found significant reductions in *Veillonellaceae* throughout the intervention period corresponding with a shift in intestinal microbial activity (Bonder et al., 2016). For extreme diets, animal models have provided an opportunity to study the impact that extreme alterations in diet have on the intestinal microbiome and demonstrate shifts in both 16S rRNA intestinal microbiome and microbial physiology. For example, mice fed a diet high in fat for a 12-week period experienced significant weight gain corresponding with a significant increase in *Rikenellaceae*, a microbe recently found to be associated with the development of type II diabetes (Daniel et al., 2014; Qin et al., 2012). Similarly, the consumption of a high-fat, high-sugar diet within a humanized mouse model resulted in a notable increase in adiposity that was successfully transplantable and replicable via fecal material transplant to recipient gnotobiotic mice (Turnbaugh & Gordon, 2009). The aforementioned studies help to support the theory that intestinal microbial composition can be influenced by diet.

Despite one's existing intestinal microbes, rapid shifts in the intestinal microbiome occur in response to alterations in diet, further demonstrating the ability of diets to regulate intestinal

microbiome composition on a day-to-day basis. Analysis of the 16S rRNA intestinal microbiome of individuals who alternated between animal- and plant-based diets observed a significant shift in microbes within 24 hours after a different diet was introduced (David et al., 2014). Most notably, the animal-based diet induced a significant increase in colonization by bile-tolerant microorganisms i.e., *Bacteroides* and decreased colonization by microbes that metabolize plant polysaccharides i.e., *Firmicutes* within 24 hours (David et al., 2014). Similar changes in the microbiome were observed within 24 hours of individuals shifting from a high-fat, low-fiber diet to a low-fat, high-fiber diet (Wu et al., 2011) over a 10-day period.

The ability of one's diet to selectively and rapidly alter intestinal microbiome composition is an important consideration in host health and intestinal microbiome composition. Because the baseline adult intestinal ecosystem is established by three years of age, periodic fluctuations in the diets of healthy adults are unlikely to significantly hinder overall colonization resistance of the gut to the point of causing disease. However, long-term trends in microbiome colonization associated with specific diets demonstrate how diet uniquely shapes the 16S rRNA intestinal microbiome and thus the potential for colonization resistance. Because of this, the current FMT donor screening process recommends deferral of individuals who choose to adopt diets that are considered extreme, such as gluten-free or high-protein based diets, regardless of the rationale. However, FMT donors that follow a balanced and traditional diet, such as animal- or plant-based diets, are allowed to donate fecal material. In spite of the role of diet in shaping the intestinal microbiome, little long-term data exist to support whether any particular diet is associated with optimal donor host health and passing the donor screening process. This study postulated that such data would be valuable to help determine whether certain FMT donors, based on diet, were more likely to pass the donor screening.

Alcohol Use. Similar to host diet, the use of alcohol has been shown to influence the intestinal microbiome. However, unlike research into the effects of diet on the intestinal microbiome, research on alcohol use and the composition of the intestinal microbiome is limited and has been primarily conducted in animal models with a focus on advanced human disease states such as alcoholic liver disease and cirrhosis. Within the available research literature there is a general consensus that chronic alcohol abuse results in a significant increase in bacterial overgrowth throughout the gastrointestinal tract (Bode, Bode, Heidelbach, Durr, & Martini, 1984). The degree of this overgrowth has been found to correlate with the presence of alcohol in the gut and worsens with increasing severity of alcohol-induced disease (P. Chen & Schnabl, 2014). Because certain intestinal microbes can metabolize alcohol, individuals who chronically abuse alcohol tend to experience intestinal dysbiosis marked by a significant increase in *Proteobacteria* and decrease in *Bacteroidetes* when compared to a healthy human microbiome (Y. Chen et al., 2011; Mutlu et al., 2012). While the underlying mechanism between intestinal dysbiosis and pathogenic alcohol-induced disease remains unclear, trials on the use of probiotics in individuals with alcohol-induced disease have demonstrated reductions in disease symptoms suggesting that the intestinal microbiome is likely associated with disease expression to some degree.

FMT donor screening programs that exclude donors based on alcohol consumption due to health related reasons generally follow guidelines provided by U.S. Dietary Guidelines (Centers for Disease Control and Prevention, 2017c). Based on these guidelines, donors who report regular weekly consumption of alcohol greater than or equal to seven drinks per week for women or fourteen drinks for men are excluded from donating. Although research on the impact of alcohol on the 16S rRNA intestinal microbiome has not been conducted in disease-free healthy

adults who consume low to moderate levels of alcohol, the intestinal microbiome has been shown to rapidly evolve based on the nutrients made available through dietary consumption. Low to moderate weekly consumption of alcohol is therefore likely to have an influence on the intestinal microbiome much like that of other dietary or nutrient intake in otherwise healthy FMT donors. As such, varying levels of alcohol intake may result in dysbiosis or altered levels of intestinal resident microbes. Based on this information, this study postulated that further exploration of the relationship between alcohol use and donor passage of the FMT screening program was warranted as it may subsequently influence the microbiome diversity of donor fecal material and clinical efficacy of FMT.

Tobacco Use. The impact of tobacco use on individual health has been well documented over several decades of research. Long-term tobacco use has been shown to negatively impact nearly every part of the human body, increasing one's risk for chronic disease, susceptibility to cancers, disease morbidity, and mortality (Centers for Disease Control and Prevention, 2017c). Regardless of the delivery route, the use of tobacco has also been asserted to increase susceptibility to certain gastrointestinal related disorders, such as Crohn's disease (X. C. Morgan et al., 2012) and colorectal cancers (Giovannucci, 2001; Giovannucci & Martinez, 1996) and places one at increased risk for tobacco related medical complications. In a recent study on the impact of smoking on patients with IBD, researchers found that current smokers with Crohn's diseases were at increased risk of requiring surgery and former smokers with ulcerative colitis were more likely to undergo a colectomy than never smokers (Kuenzig et al., 2016).

Tobacco use has also been shown to induce changes in the intestinal microbiome. In a study on the impact of tobacco use in individuals with IBD, researchers found that the abundance of a specific bacteria from the *Firmicutes* phylum, *Anaerostipes*, was significantly decreased by

greater than 60% in current and former smokers (Cosnes, 2004). This finding was further supported in a more recent study looking at the evolution of the intestinal microbiome after smoking cessation, which demonstrated that reduced intestinal levels of *Firmicutes* and *Actinobacteria* rebounded upon complete cessation of smoking (Biedermann et al., 2013). The significant increase in diversity in the intestinal microbiome of these individuals occurred within four weeks after smoking cessation and was maintained at the eight-week follow-up visit. Most notably, the use of tobacco appears to significantly impact butyrate producing microbiota (Biedermann et al., 2013; Cosnes, 2004; Nagao-Kitamoto et al., 2016), which are important in overall intestinal health.

Tobacco's impact on the intestinal microbiome is not surprising given the existing research findings on tobacco use and multisystem disease. As such, for programs that incorporate tobacco use during their screening, FMT donors are asked to quantify their past and current smoking history during the in-person screening and are classified according to the terminology set forth by the Centers for Disease Control (CDC) (Centers for Disease Control and Prevention, 2017b). Donors who meet the definition of daily smokers as defined by the CDC as one cigarette or more per day are typically deferred from donating fecal material. However, clear deferral exclusions do not exist for FMT donors who smoke less than one cigarette per day or former smokers, so these donors may be allowed to donate fecal material. While a recent study was unable to find a difference between the 16S rRNA intestinal microbiome of non-smokers and smokers (Tu et al., 2017), the study did not distinguish between current and former smokers who were grouped together categorically. Further, the Biedermann study (2013) demonstrated that the intestinal microbiome is capable of effectively and rapidly restoring a diverse intestinal ecosystem upon complete cessation of tobacco use, suggesting a difference may exist between

former smokers and current or intermittent smokers who, for example, smoke infrequently or during social events alone. Despite the plasticity of the intestinal microbiome, little information exists about the impact of infrequent or intermittent use of tobacco on the intestinal microbiome. As such, this study postulated that further evaluation of the impact of varying degrees of tobacco use on FMT deferral rates could provide valuable information to improve the FMT screening process and clinical efficacy rates for donor fecal material.

Seasonality. In infectious disease medicine, seasonality refers to a "periodic surge in disease incidence" that corresponds with different seasons across one full calendar year (Fisman, 2007). Although the rationale is not entirely understood, seasonal changes have been observed for both respiratory and gastrointestinal infectious diseases. In particular, the incidence of respiratory illnesses, such as rhinovirus, tends to increase with the resumption of the fall academic school year and remain elevated through early spring (Monto, 2002). The emergence of influenza tends to be slightly delayed with cases occurring more often during the winter and spring seasons (Monto, 2002). Certain gastrointestinal related infectious diseases, such as rotavirus, also tend to follow a similar pattern with a peak incidence observed in the U.S. during the winter months (Cook, Glass, LeBaron, & Ho, 1990). Because the risk reduction approach aims to defer individuals with an increased risk for transmitting infectious and communicable diseases, the rate of donor deferral may vary significantly from one season to the next. This challenge is further complicated due to asymptomatic carriage, which may be higher during periods of peak incidence (Pickering, Bartlett, Reves, & Morrow, 1988). As such, an unanticipated variable in the donor screening process may be the varying rates of donor deferral during different seasons due to fluctuations in infectious disease occurrence and risk.

In order to accurately gauge the impact of seasonality on the intestinal microbiome, researchers must control for many variables that may confound the results. For example, one's diet naturally changes according to the nutrients available during different seasons. Because diet has been shown to have a notable and rapid impact on the intestinal microbiome, it can be difficult for researchers to corroborate changes in the intestinal microbiome to seasonality alone and thus few published studies exist. However, one study conducted on 60 members of an indigenous group attempted to associate changes in the intestinal microbiome with the different seasons. While a distinct and clear association among all phylum levels of the 16S rRNA intestinal microbiome was observed with seasonality, the researchers were admittedly not able to entirely control for the effects of diet, which differed markedly between seasons based on the availability of fresh produce (Davenport et al., 2014). A similar study conducted in Japan also observed a change in microbiome composition with seasonality, specifically *Bifidobacterium* and *Lactobacillus*. Interestingly, the researchers suggested that the significant change in *Bifidobacteria* could be associated with changes in fermented milk consumption during different seasons (Hisada, Endoh, & Kuriki, 2015). However, the researchers were not able to attribute the change in *Lactobacillus* with diet or fermented milk consumption, suggesting that seasonal variation may have contributed to this difference.

While research has been done to evaluate the impact of infectious gastroenteritis on the intestinal microbiome, there is little information available regarding acute respiratory diseases and their impact on the intestinal microbiome. Because of the risk reduction approach to infectious disease, FMT donors who exhibit active symptoms for any infectious disease are deferred from donating fecal material. The FMT donor screening further reduces this risk by excluding donors found to be asymptomatically affected with infectious diseases during the

laboratory screening. Despite these efforts, several theories purport the impact of seasonality on host health and the intestinal microbiome independent of infectious disease prevalence, which may be predictive of passing the FMT donor screen and microbiome composition. These include changes in atmospheric conditions, different light/dark periods, and cyclical variations in host immune function (Dowell, 2001). While seasonality may be found to contribute to the cyclical appearance of infectious disease outbreaks, because of the bidirectional relationship between the host immune system and intestinal microbiome, it is possible seasonality may also potentially skew intestinal colonization resistance. As such, this study postulated that evaluating for a seasonal variation in FMT donor deferral could allow stool banks and providers to proactively plan in advance for potential FMT donor shortages or target recruitment to ensure the collection of fecal material with more optimal intestinal microbiome composition diversity.

Microbiome Features.

Fecal Microbiome Diversity. The value of understanding how the intestinal microbiome evolves and maintains a mature homeostatic and symbiotic ecosystem becomes evident when consideration is given to the impact that the microbiome has in shaping the intestinal ecosystem, influencing microbial inhabitants, and supporting host defenses against infectious disease. When significant alterations to the natural ecosystem occur, intestinal homeostasis and immune expression can be negatively impacted and result in diseases such as those caused by pathogenic *Clostridium difficile* bacteria. However, systematic review of the literature has demonstrated FMT clinical efficacy rates of up to 94% in the treatment of CDI recurrence (Kassam et al., 2013; Sha et al., 2014) suggesting that re-colonization of the perturbed intestinal microbiome with healthy donor fecal material can prevent disease recurrence.

Prior to the emergence of stool banks, FMT donors were utilized on a once or twice basis making it difficult to meaningfully study the critical components of the intestinal microbiome in treating recurrent CDI. As such, most research conducted on FMT donor 16S rRNA intestinal microbiome composition has focused on clinical efficacy rates in the prevention of CDI recurrence. Research has begun to emerge evaluating FMT donor characteristics and 16S rRNA intestinal microbiome composition with clinical efficacy in preventing CDI recurrence. In 2016, researchers evaluating the clinical efficacy rates of four FMT donors in an enema treatment of 34 cases of recurrent CDI did not find a significant difference in the FMT clinical efficacy rates among the donors (Ray & Jones, 2016). However, neither FMT donor characteristics nor their 16S rRNA intestinal microbiome composition were evaluated as part of this study. More recently, a pilot study on 35 FMT donors looking at the impact of diet, alcohol, and tobacco intake on 16S rRNA intestinal microbiome composition and FMT clinical efficacy was inconclusive (Tu et al., 2017). However, the finding may be attributed to limitations in the study design, such as the use of overly inclusive definitions for predictor variables and a delayed time span between the completion of the food frequency questionnaire and collection of the fecal sample for 16S rRNA microbiome evaluation. Interestingly, a recent study by Budree and colleagues (2017) compared the 16S rRNA intestinal microbiome of FMT donors with superior clinical efficacy rates (greater than 90%) with FMT donors demonstrating anticipated or normal clinical efficacy rates of 80% to less than 90% and found that 16S rRNA intestinal microbiome diversity did not differ significantly between the groups. No additional FMT donor characteristics were considered or analyzed as part of that study.

Summary

Prospective FMT donor selection and screening continue to vary significantly across institutions (Kump, Krause, Allerberger, & Hogenauer, 2014). The variability observed in prospective FMT donor screening programs is most notable with the direct donor approach. The first public stool bank opened in 2013 has screened over 1,000 FMT donors and provides clinician access to FMT material from multiple anonymous donors. Despite the availability of this donor information, little research has been conducted evaluating donor factors that may improve passage of the FMT donor screening program. As such, few improvements have been made to the prospective FMT donor screening process, which remains logistically challenging for both stool banks and individual physicians and hospitals that continue to use the directed donor approach.

To date, no specific microbes or 16S rRNA based intestinal microbiome levels of diversity have been correlated with consistently providing optimal clinical efficacy in preventing CDI recurrence. Further, research looking at 16S rRNA intestinal microbiome composition and FMT clinical efficacy has been limited. Intestinal ecosystem diversity is a critical component required for the sustainability of colonization resistance and ecosystem stability in patients with recurrent CDI. Knowledge of the 16S rRNA microbiome composition and level of diversity of optimal FMT donors could serve to validate the 16S rRNA microbiome-based biomarker and improve prospective donor screening and FMT clinical cure.

While a relatively new occurrence, the emergence of public stool banks has eased some of the logistical barriers to the preparation of FMT material and introduced greater consistency in prospective FMT donor screening and stool processing. However, in a survey of infectious disease physicians, the complexity and cost of donor screening was cited as the second most

common reason for not offering FMT (Bakken, Polgreen, Beekmann, Riedo, & Streit, 2013). The barriers to FMT have driven some patients to search for alternative options including unsupervised at home self-administration utilizing unscreened donor fecal material that could carry infectious or microbiome mediated diseases (Smith, Kelly, & Alm, 2014). As such, there is a clear need for improvement in the prospective FMT donor screening process and clear guidelines on the optimal intestinal microbiome composition for prospective FMT donors.

Chapter 3. Study Design and Methodological Approach

Overview of Study Design

This study was conducted as a secondary analysis of data collected from a cohort of fecal microbiota transplant (FMT) donors screened by OpenBiome, a public stool bank. The hypothesis for Aim 1 examined the association between previously unstudied donor characteristics and environmental influences on the donor's likelihood of passing the company's current screening regimen. That regimen categorized prospective donors as "passed" or "deferred" through a multistep process that included prescreening, clinical, and laboratory analysis phases. This study used logistic regression to determine whether additional donor host (gender, age, body mass index [BMI], and frequency of bowel movements) and environmental (diet, alcohol use, tobacco use, and seasonality based on screen date) factors affected the likelihood of passing the screening. Donors who passed the screening and provided stool for transplant into a patient had the microbial diversity of their stool quantified, but the relationship between that diversity and the clinical efficacy of the procedure has not been studied. As such, the hypothesis for Aim 2 used a simple linear regression to determine whether the microbial diversity of the donor stool predicted the rate of clinical cure.

Hypotheses

1. There is a statistically significant association between one or more characteristics of prospective FMT donors i.e., donor gender, age, BMI, frequency of bowel movements, diet, alcohol use, tobacco use, or seasonality based on screen date (predictor variables) and passing the FMT donor screening program (outcome variable).

2. There is a statistically significant association between the intestinal microbiome alpha diversity of active FMT donors (predictor variable) and *Clostridium difficile* infection (CDI) clinical efficacy (outcome variable).

Site and Sampling

This study was conducted as a secondary analysis of FMT donor clinical and microbiome related data collected at OpenBiome, a not-for-profit stool bank located in the United States (U.S.) in Somerville, Massachusetts. The town of Somerville is located within the Greater Boston Metropolitan area. OpenBiome recruited and collected FMT donor data between August 2013 and December 2016 under a Massachusetts Institute of Technology (MIT) Human Subjects approved research protocol titled "Building a repository of healthy stool samples for fecal microbiota transplantation." The aim of the study was to facilitate the collection of healthy donor stool for use as FMT material to prevent recurrent CDI and to promote research on the human microbiome. FMT donors who enrolled in the study were over 18 years of age and lived within close proximity of less than one hour's travel distance to the stool bank donation location. For the purpose of testing the hypotheses presented by this research proposal, data for all FMT donors recruited between August 2013 and December 2016 were assessed for inclusion in this study.

The following sections provide details regarding the operating procedures for OpenBiome that were used to collect the donor data to be utilized by this study. This study was conducted as a retrospective analysis of existing data already collected and stored in OpenBiome's data repositories. As such, all laboratory testing, stool sample processing, and 16S ribosomal ribonucleic acid (rRNA) intestinal microbiome analyses were conducted under

OpenBiome's approved research protocol. No additional laboratory or sample processing were required or conducted as part of this research study.

Stool Bank Operating Procedures. Standard operating procedures for OpenBiome's donor screening program included interacting with prospective FMT donors beginning at initial online contact during the pre-screening, in-person prospective FMT screening which included nursing and physician clinical assessment and laboratory based testing, and ending with final collection of stool samples for 16S rRNA intestinal microbiome analysis. Data generated during these procedures were stored in data repositories managed and controlled by OpenBiome.

Pre-Screening. Prospective FMT donors were recruited through flyers and by word of mouth. Individuals who expressed interest in donating their stool were directed to register and complete the online pre-screening survey. The survey included questions about demographics, logistics, and health status. Health related questions were extracted directly from the Stool Donor Health Questionnaire (DHQ), the instrument used in the more comprehensive clinical assessment phase of the screening process described below. The survey took approximately ten minutes to complete online. Once submitted, a trained nurse and a physician study staff member as needed reviewed each prospective FMT donor's answers to the questions to determine whether or not the prospective FMT donor met the inclusion criteria and could continue on to the in-person clinical assessment portion of the FMT donor screening program. Prospective FMT donors who met inclusion criteria were contacted to schedule an appointment for a clinical assessment and submit biological samples for laboratory screening.

Clinical Assessment. The aims of the clinical assessment and laboratory screens were to evaluate the prospective FMT donor's health, behavioral, and social history for presupposed risks for infectious diseases and hypothesized microbiome-mediated health conditions that may

be transmissible via stool. In-person clinical assessments were conducted by a nurse or physician study staff member and were overseen by the study's Internal Medicine Physician Investigator.

Prior to participating, the study staff member reviewed the purpose of the study in detail with the prospective FMT donor and answered any questions about participation. If the prospective FMT donor was interested in participating, informed consent utilizing OpenBiome's MIT Human Subjects Approved consent form was obtained. The prospective FMT donor and the study staff member signed two copies of the consent form, one copy was provided to the prospective FMT donor and the second retained at OpenBiome.

During the clinical assessment, prospective FMT donors were asked to complete OpenBiome's full Stool DHQ in-person while at the stool bank. The Stool DHQ was developed by a clinical advisory board of twelve thought leaders and industry experts in infectious disease and gastroenterology with additional guidance from the U.S. Food and Drug Administration (FDA). Questions were based on the U.S. FDA recognized full-length American Association of Blood Bank's DHQ. The Stool DHQ consists of over 100 closed- and open-ended questions aimed at obtaining a comprehensive health, behavioral, and social assessment of the prospective FMT donor's risk factors for infectious or microbiome mediated diseases.

Once completed, answers on the Stool DHQ were reviewed with the prospective FMT donor by a nurse or physician during the clinical interview. When needed, additional inquiry was requested to ensure accuracy of the response and clarify any uncertainty in the reported answers. As part of the clinical assessment, relevant vital signs were obtained by a nurse or physician including height, weight, and BMI. The entire in-person clinical assessment lasted approximately 60 minutes. If any apparent risk factors were identified from answers on the Stool DHQ, prospective FMT donors were deferred temporarily or permanently from donating. Prospective

FMT donors with no apparent risks for infectious or communicable diseases were asked to submit a stool sample for pathology testing and blood sample for laboratory based serological testing.

Laboratory Screening. Stool and serological laboratory tests were selected from the standard screening guidelines utilized by Blood Banks and further recommended by OpenBiome's clinical advisory board in collaboration with feedback from the FDA. Prospective FMT donors were provided with a stool collection kit and received training in safe and sterile stool sample collection techniques. Prospective FMT donors were asked to collect a stool sample at home or provide a sample on site and deliver a fresh stool sample within less than one hour of passage to the stool bank for processing. Upon receipt, OpenBiome's laboratory technicians placed aliquots of the prospective FMT donor's stool sample into the proper collection containers and sent the samples to a Clinical Laboratory Improvement Amendments (CLIA) approved laboratory for comprehensive screening for stool transmissible pathogens.

Additionally, a nurse or physician provided prospective FMT donors with a laboratory requisition for serological based screening to assess general prospective FMT donor health and to screen for transmissible infectious diseases. Prospective FMT donors were then provided with a list of local CLIA approved laboratories and asked to take the requisition to have their blood drawn for testing. Trained phlebotomist professionals employed by the CLIA approved laboratory conducted the phlebotomy and blood collection procedures and submitted the samples to the CLIA approved laboratories for testing.

Upon receipt of the stool and serological based test results, a nurse and physician reviewed all laboratory results to assess whether the prospective FMT donor met inclusion criteria. Results were subsequently communicated with the prospective FMT donors. Any

abnormal result was deemed an exclusion criteria and the prospective FMT donor was either temporarily or permanently deferred from further participation in the study.

Stool Collection and Preparation. Donors who met all inclusion criteria after completing the prospective FMT donor screening process were enrolled into OpenBiome's stool donation program as active FMT donors. Active FMT donors were provided with stool collection kits and asked to donate stool samples for use as FMT transplant material for clinical treatment to prevent recurrent CDI. In return, active FMT donors received remuneration of $40 per accepted stool donation. Institutions that requested and utilized active FMT donor material voluntarily returned de-identified clinical efficacy data to the stool bank utilizing OpenBiome's FMT Follow-up Form.

During the donation period, an active FMT donor's stool sample was selected and utilized to conduct 16S rRNA intestinal microbiome sequencing. Upon delivery of the sample to the stool bank laboratory by the active FMT donor, a member of the research team processed and placed aliquots of the stool sample into microtubules for freezer storage at -80° Celsius. Because concurrent sample sequencing reduces the risk for systematic bias between runs, stool samples from multiple active FMT donors were stored until a sufficient number of stool samples were collected. Once this amount was reached, the batch of active FMT donor stool samples was sent to a laboratory specializing in conducting intestinal microbiome molecular characterization using a 16S rRNA microbiome platform.

16S rRNA Microbiome Sequencing. Stools samples collected for microbiome analysis underwent standard 16S rRNA intestinal microbiome genetic sequencing. Power Mag Microbiome kit with glass beads was used for repeated bead-beating cycles. DNA extracts that were not used immediately were stored at -20°C until further use. DNA extracts were then

amplified targeting the V4 region of the bacterial 16S ribosomal gene (Caporaso et al., 2011; Kozich, Westcott, Baxter, Highlander, & Schloss, 2013). Resulting amplicons were cleaned and quantified, pooled equimolarly, and then paired end sequenced on the Illumina MiSeq platform according to manufacturer's instructions.

Prior to data analysis, raw sequencing data were filtered and processed according to the UPARSE operational taxonomic unit (OTU) pipeline to create sequence clusters at various identity thresholds. Taxonomic identity assignment was based on review of the Greengenes reference database at a clustering at 97% identity or greater and RDP classifier. Results of the donor stool microbiome sequencing were stored as raw reads and processed taxonomic data in OpenBiome's Microbiome Data Repository.

Stool Bank Inclusion and Exclusion Criteria. The following OpenBiome stool bank determined inclusion and exclusion criteria were applied during the pre-screening and clinical assessment stages of the prospective FMT donor screening program. OpenBiome utilized a step approach to evaluating study inclusion due to the effort and costs associated with conducting in-person clinical assessments and laboratory screenings. Only upon passing all inclusion criteria during the pre-screening and clinical assessment stages was the prospective FMT donor allowed to actively donate stool for FMT transplantation. As such, 16S rRNA intestinal microbiome and clinical efficacy data are available only for active FMT donors who passed all pre-donation study inclusion criteria.

Pre-Screening. Adults 18 years of age or older with access to the Internet were directed to complete the Stool Bank Registry online. Prospective FMT donors were deferred during the pre-screening process if their answers included any of the following exclusion criteria:

- Under 18 years of age or greater than 50 years of age
- BMI of 30 or greater

- Tobacco use of three cigarettes per week or greater
- Use of antibiotics, antifungals, or antivirals within the past three months
- Vaccination with a live attenuated virus within the past two months
- Travel within the past 12 months to an International SOS (ISOS) identified high risk country
- Unable or unwilling to travel to the stool bank facility
- Unable or unwilling to donate stool samples in person 4 times or more per week
- Earning cash as the sole motivation for donating

Clinical Assessment. Prospective FMT donors that underwent an in-person clinical assessment and laboratory screening were temporarily or permanently deferred from the study if they met any of the exclusion criteria specified on the Stool DHQ. In addition to reassessing the pre-screening exclusion criteria, prospective FMT donors whose daily commute to the stool bank was determined to be greater than one hour were excluded from participating. The one-hour time limit was imposed by OpenBiome to ensure the integrity of the intestinal microbiome in the stool donation. For proprietary reasons and to protect the continued integrity of OpenBiome's Stool DHQ, broad exclusion categories have been provided. Prospective FMT donors with a remarkable health or medical history in any of the following areas were deferred either temporarily or permanently from donating:

- Atopy, asthma, or allergies
- Autoimmune conditions
- Cancer history
- Cardiovascular and Metabolic conditions
- Current health status
- Diet and exercise
- Family history
- Gastrointestinal conditions
- Infectious disease history
- Medications and supplements
- Mental health and well-being
- Musculoskeletal conditions
- Neurological conditions
- Sexual behavior and history
- Social history
- Surgical and other medical history

- Travel history
- Women's health
- Vital signs

Laboratory Screening. Upon completion and passing of the Stool DHQ and in-person clinical assessment, prospective FMT donors were asked to submit biological samples for serological and stool-based testing for infectious and communicable diseases. Complete blood count and liver function tests were also conducted to assess overall prospective FMT donor health status. Prospective FMT donors that received a positive infectious or communicable disease result or abnormal result for any of the serological laboratory tests listed below were deferred either temporary or permanently from donating fecal material:

- Complete blood count
- Liver function panel
- Adenovirus
- *Campylobacter*
- Carbapenemase producing gram-negative rods
- *Clostridium difficile*
- Cryptosporidium
- *Escherichia coli* O157
- Extended spectrum beta-lactamase (ESBL) producing organisms
- Giardia with reflex to ova and parasites
- *Helicobacter pylori*
- Hepatitis A, B, and C
- Human immunodeficiency virus antibodies types 1 and 2
- Human T-cell lymphotropic viruses types 1 and 2
- Isospora and Cyclospora
- *Listeria monocytogenes*
- Methicillin-resistant *Staphylococcus aureus*
- Norovirus
- Rotavirus
- *Salmonella*
- *Shigella*
- *Treponema pallidum*
- Vancomycin-resistant *Enterococci*
- *Vibrio*

Stool Bank Data Repositories. The data for this study were extracted from three different OpenBiome stool bank data repositories including the Stool Donor Registry, Clinical Assessment and Safety (CAS) Database, and Microbiome Data Repository.

Stool Donor Registry. The Stool Donor Registry was created to assist during the online pre-screening of prospective FMT donors. This data repository contains information on healthy adults over 18 years of age who expressed interest in donating fecal material to OpenBiome and who completed the online survey. The registry contains data for over 8,000 prospective FMT donors including limited biodemographic information, motivation and logistics, and answers to a subset of questions taken directly from the Stool DHQ. Donor pass or deferral status from the pre-screening is also noted for each prospective FMT donor included in the database.

Clinical Assessment and Safety Database. The Clinical Assessment and Safety (CAS) Database consists of data from a subset of prospective FMT donors that were recruited from the Stool Donor Registry. These prospective FMT donors met study inclusion criteria during the pre-screening step and travelled to OpenBiome where they underwent an in-person clinical assessment and laboratory screening i.e., stool pathogen and serological testing, to verify eligibility and safety of their stool samples. The database contains clinical data for over 700 prospective FMT donors including answers to the entire Stool DHQ and results of prospective FMT donor stool pathology and serological laboratory screening tests conducted on prospective FMT donors to determine eligibility to donate fecal material for clinical use. Donor pass or deferral status at each step of the clinical screening process is noted for every prospective FMT donor. Additionally, clinical efficacy rates of active FMT donor fecal material in preventing CDI recurrence were recorded in the database for active FMT donor fecal material utilized as part of clinical care fecal transplants.

Microbiome Data Repository. The Microbiome Data Repository contains raw 16S rRNA taxonomic microbiome data from active FMT donors who passed the clinical screening and donated stool samples to OpenBiome. The database contains 16S rRNA intestinal microbiome sequencing data for over 50 donors. This information is stored separately due to the large data storage capacity required to store the microbiome data files.

Study Sample Size

Based on power analysis performed using G*power© statistical analysis software, a minimum of 568 prospective FMT donors was required for the multiple variable logistic regression analysis (Aim 1) to reduce the chance of type I and type II statistical errors and to attain adequate power (0.80). Prospective FMT donor data required for the multiple variable logistic regression portion of this study was collected on all prospective FMT donors in the CAS Database who were recruited between August 2013 and December 2016, totaling over 750 prospective FMT donors. Although prospective FMT donor recruitment by OpenBiome continued after December 2016, the donor population recruited after this date was performed under a clinical study approved by a different ethics board. The prospective FMT donor population recruited through December 2016 was deemed sufficient to that ensure power was achieved and to account for any unexpected challenges with data integrity.

Intestinal 16S rRNA microbiome and clinical efficacy data were collected for a subset of active FMT donors who passed OpenBiome's prospective FMT donor screening program and donated fecal material for clinical use. All active FMT donors with available microbiome 16S rRNA and clinical efficacy data were evaluated for inclusion in the analyses and dependent on the availability of active FMT donor data.

Study Variables

Aim 1. **Identify additional, previously unstudied characteristics of prospective FMT donors that are predictive of passing a public stool bank's current screening process**. A multiple variable logistic regression model was utilized as part of Aim 1 to test the hypothesis that there is a statistically significant association between one or more characteristics of prospective FMT donor characteristics (predictor variables) and passing the FMT donor screening program (outcome variable). Theoretical and operational definitions for the prospective FMT donor characteristics tested as part of the statistical analysis portion of the predictive analysis are provided in Table 1.

Table 1.
Donor characteristics of interest for prospective FMT donors

Outcome Variable	**Theoretical Definition**	**Operational Definitions**
FMT Donor Status (dichotomous)	• Final determination of prospective FMT donor's eligibility to donate stool samples based on results from the clinical screening process.	• The prospective FMT donor's eligibility to donate stool will be assigned to one of the following values: o 0= defer (Donor not able to donate temporarily or permanently) o 1= pass (Donor able to donate immediately)
Predictor Variable	**Theoretical Definition**	**Operational Definitions**
Gender (dichotomous)	• Physiological sex of FMT donor	• Confirmed through donor self-report and assigned to one of the following values: o 1= male o 2= female
Age (discrete)	• Age of FMT donor at the time of study participation	• Recorded as a discrete numerical value of years that the donor self-reports he/she has been living since birth
Body Mass Index (continuous)	• A numerical estimate of body fat composition	• Donor weight measured with a digital scale and recorded by the stool bank clinician in kilograms divided by the square of the body height measured by the clinician with a height rod in meters. Universally written as a three-digit number rounded to the first decimal and expressed in terms of kg/m^2

Table 1 (continued)

Predictor Variable	Theoretical Definition	Operational Definitions
Frequency of bowel movements (discrete)	• An estimate of the average number of bowel movements or stools a person passes per day	• The average number of bowel movements a donor self-reports as passing each day, rounded to the first decimal.
Diet (categorical)	• A specific course of food to which one restricts oneself, either to lose weight or for medical or ethical reasons	• Type of diet regularly followed as self-reported by the donors and categorized as follows: o 1= Diets that include meats o 2= Diets that exclude animal-based meats i.e., vegetarian, pescatarian, vegan o 3= Diet not specified
Alcohol Use (continuous)	• An estimate of the current alcohol intake on a weekly basis, measured by the number of servings equal to 0.6 fluid ounces or 14 grams of alcohol	• An average of the prospective FMT donor's self-reported weekly consumption of alcoholic beverages, rounded to the first decimal place.
Tobacco Use (categorical)	• Estimate of one's historical and current direct (excluding secondary) exposure to tobacco through cigars or cigarettes, pipe smoking, or other tobacco-based products such as smokeless chewing tobacco, hookah, kreteks, e-cigarette, vaping, etc.	• Prospective FMT donor's self-reported history of direct exposure only to tobacco based on one of the following categories: o 1= Never smoker (An adult who has never smoked, or who has smoked less than 100 cigarettes in his or her lifetime) o 2= Former smoker (An adult who has smoked at least 100 cigarettes in his or her lifetime but who had quit smoking at the time of interview) o 3= Current Smoker (An adult who smokes now but does not smoke every week)
Seasonality (categorical)	• Clinical assessment date conducted within the four meteorological seasonal cycles contained within one calendar year	• Date of the prospective FMT donor's clinical assessment categorized into one of the following groups: o 1= Winter (December 1 to February 28 or 29 in leap years) o 2= Spring (March 1 to May31) o 3= Summer (June 1 to August 31) o 4= Fall (September 1 to November 30)

Outcome Variable. The primary outcome of interest for the predictive multiple variable logistic regression was FMT donor status defined as pass or deferral. FMT donor deferral is defined as the non-acceptance of a prospective FMT donor upon conclusion of the prospective FMT donor screening process. Prospective FMT donors who pass all screens and are approved to donate fecal material for clinical treatments are referred to as active donors within the transplant industry. This study utilized the FMT donor status outcome variable to develop a model that identified previously unstudied FMT donor characteristics that are most predictive of becoming active donors upon completion of the prospective FMT donor screening process.

Predictor Variables. The choice of variables utilized in the predictive analysis were guided by the conceptual framework. The selected predictor variables were considered influential factors in health and disease of prospective FMT donors and supported by the research literature as potentially associated with variations in the intestinal microbiome and colonization resistance. Specific prospective FMT donor characteristics that were of concern in this study were FMT donor host factors including gender, age, BMI, and frequency of bowel movements, and environmental factors including diet, alcohol and tobacco use, and seasonality based on screen date. These predictor variables were included to assess the association and significance of prospective FMT donor characteristics with passing the FMT donor screening.

Aim 2. **Determine whether the microbial diversity of active FMT donor stool, as measured by 16S rRNA, is related to FMT clinical efficacy, as measured by the recipient's clinical cure**. A simple linear regression analysis was utilized to test the hypothesis that there is a statistically significant association between active FMT donor intestinal microbiome 16S rRNA based alpha diversity (predictor variable) and FMT clinical efficacy (outcome variable).

Theoretical and operational definitions for the variables tested as part of this statistical analysis are provided in Table 2.

Table 2.
Donor characteristics of interest for prospective FMT donors

Outcome Variable	Theoretical Definition	Operational Definitions
Clinical efficacy (continuous)	• The proportion of cases classified as cured by the donor's fecal material	• The total number of cases reported as cured divided by the total number of cases and reported as a whole proportion.
Predictor Variable	**Theoretical Definition**	**Operational Definitions**
Alpha Diversity (continuous)	• A numerical measure of richness and evenness of the microbiota in a specific ecosystem based on the presence of bacterial 16S rRNA, a.k.a. Shannon's diversity index.	• The total sum of bacteria after dividing the number of individual species found in a sample by the number of all species, multiplied by its natural log. Both microbial richness and evenness of the community increase Shannon's diversity index.

Outcome Variable. The primary outcome of interest utilized in the simple linear regression analysis for Aim 2 was the clinical efficacy of FMT in CDI patient recipients of FMT donor fecal material. Clinical efficacy rates were collected from the FMT donor database. The clinical efficacy rate is a proportional value calculated for each active FMT donor and represents the percentage of clinical cure. The clinical efficacy rate was determined by dividing the total number of recurrent CDI cases identified as clinically cured by the total number of cases where the active FMT donor's material was used. Clinical efficacy data was provided by treating physicians and collected prior to or during the eight-week follow-up visit.

Predictor Variable. The primary predictive variable considered in the simple linear regression analysis was 16S rRNA based intestinal microbiome alpha diversity as represented by the Shannon diversity index. Alpha diversity provides an estimate of the mean phylogenetic diversity found within an active FMT donor's intestinal ecosystem (Tyler et al., 2014). This

mean estimate is calculated by using multiple reads of an active FMT donor's fecal sample to identify the number of different 16S rRNA sequences within a specific fecal sample and is represented by the Shannon diversity index measure.

Data Collection Procedures and Strategies

A data collection tool was used to record information from the data repositories about specific biodemographic, behavioral, and physiological variables to be used as part of this study. The data collection tool was created and managed in REDcap, a research electronic data capture tool and a Microsoft Excel© spreadsheet for upload into the statistical software program, IBM® Statistical Package for the Social Sciences (SPSS), where the data were analyzed upon completion of data collection. On rare occasions where data about a variable were not readily available within the data repositories, the data were imputed through a review of the source documents. Access to and collection of data was done in a de-identified manner using the anonymous derived donor identification number. No FMT donor identifiers were viewed or recorded as part of this study.

Data Analysis Plan

Data were entered into an IBM® SPSS statistical software program for analysis purposes. A significance level (*p* value) of 0.05 (5%) was used as the threshold to determine whether the independent predictor variables should be retained in the final models.

Descriptive Analyses. Descriptive statistics including, but not limited to, mean, frequencies and measures of central tendency and dispersion were used to summarize the demographics of the overall prospective FMT donor population and outcomes of OpenBiome's FMT donor screening program. Descriptive statistics were conducted using prospective FMT

donor characteristics collected during the pre-screening and clinical assessment stages, such as gender, age, and BMI, to describe the donor population and sample used in this study.

Descriptive statistics were utilized to provide an overview of the donor screening outcomes for OpenBiome's entire FMT donor screening program and at each step of the program. Prospective FMT donors were broken down by donor pass and deferral for the overall screening program and at each step in the screening i.e., pre-screening, clinical assessment, and laboratory investigation. The rationale for prospective FMT donor deferral was also available and was summarized at each step in the screening and for the program overall. Summarizing the rationale for deferral was used to provide an overview of donor outcomes and a descriptive breakdown of the most common reasons for donor deferral.

Aim 1 Statistical Analyses. A predictive modeling analysis utilizing multiple variable logistic regression was used to test the hypothesis that one or more prospective FMT donor characteristics are predictive of passing the FMT donor screening process. The dependent outcome variable was defined as a dichotomous categorical variable identifying whether the prospective FMT donor passed the screening process or was deferred from donating stool to be used as fecal transplant material.

Prior to analysis, categorical predictor variables established *a priori* were summarized into proportions. Correlations among the categorical variables were assessed through a chi-square analysis. Continuous variables were summarized using means, standard deviations, and t-tests to identify any pre-existing differences in the groups. A significance level (p value) of 0.05 (5%) was used as the threshold to determine whether the variables should be reevaluated or reconsidered for inclusion in the final multiple variable logistic regression analysis.

Statistical assumptions underlying the multiple variable logistic regression method were tested to verify that study data met the criteria required for a logistic regression analysis. Scatterplots for each individual predictor variable against the outcome variable were constructed and evaluated to provide a preliminary assessment of the relationship between the variables, identify potential outliers, assess for nonlinearity in the data, and verify independence of the error. Data points were evaluated to assess whether they were outside the +/-3.0 standard deviation cut-off and considered for exclusion as potential outliers.

Assumption of linearity was assessed through a plot of the residuals versus the predicted values. Scatterplots for studentized residuals of each predictor variable and probability P-P plots of the standardized residuals were produced to evaluate for constant variance across the model. Additionally, normality of the residuals and skewness for continuous variables were assessed through visual evaluation of the Standardized Residual histogram and assessed for the need for log transformation. Fischer's measure of skewness was calculated to assess the significance of any observed skewness in the variables. If required, recoding of the data and collapsing of categories was considered to ensure an equal distribution of the data across groups.

Multicollinearity between two variables is a concern with multiple variable regression analyses. As such, a correlation matrix of the variables, using Pearson correlation, was produced and assessed for collinearity among the variables. Linear regression analysis and collinearity diagnostics were run to further explore the correlational structure between the variables. A finding of multicollinearity among the predictor variables suggested that an interaction might be present, and warranted further evaluation and hypothesis testing of variables that may moderate one another.

Prior to conducting the analysis, the sample was assessed for missing values through frequency statistics and cross tabulation tables. Variables were assessed for inherent differences or patterns in the missing data. Sensitivity analysis on the missing data was conducted to evaluate the randomness of the missing values. An application for dealing with missing data i.e., listwise and case wise removal, was considered and performed. Descriptive statistics were run to assess the impact of data removal on the variables.

A stepwise multiple variable logistic regression with backward elimination was performed to test the predictors for significance. The Goodness of Fit test was evaluated to see how well the data fit the model. The likelihood ratio and Wald statistic were evaluated to assess how well the variables in the model predicted the baseline outcome. Using the significant predictors determined by this analysis, the data were fit to a final multiple variable logistic regression model that identified the prospective FMT donor characteristics most associated with passing the FMT donor screening program. Odds ratios for each variable identified by the multiple variable logistic regression model as significantly predictive of the outcome were calculated and summarized.

Aim 2 Statistical Analyses. Standard microbiome descriptive analyses were performed to describe and assess the composition of the intestinal microbiome of active FMT donors who pass OpenBiome's FMT donor screening program. The OTU table identified through 16S rRNA sequencing was utilized to create a taxonomy bar plot to illustrate the relative abundance for the ten most abundant bacterial taxa distributed at the phylum level. Limiting the analysis to the ten most abundant bacterial taxa allows researchers to accurately identify and capture the major microbial phyla present in the fecal material providing a more informative description of the microbiome composition (Tyler et al., 2014). The data from cases were also utilized to build a

heatmap to provide a visual distribution of the taxa by the most and least abundant bacterial at the phylum level. Distance matrices were used to measure beta diversity supported with partial mantel and multiple matrix regression with randomization to evaluate and explain variations in the data. Ordination plots, or Principal Coordinates Analysis (PCoA), were utilized to visualize the data present in the beta diversity index.

16S RNA based alpha diversity of the intestinal microbiome for each active FMT donor was calculated and represented by Shannon's diversity index. Generalized linear models were used to evaluate and explain any variations in the data if apparent. A box and whisker plot was created utilizing the Shannon's diversity index to compare the relative differences in alpha diversity among the active FMT donor characteristics found to be predictive of passing the FMT donor screening process.

A simple linear regression model was performed to test the hypothesis that active FMT donor 16S rRNA intestinal microbiome alpha diversity is significantly associated with FMT clinical efficacy. Prior to analysis, statistical assumptions underlying the simple regression model were tested to verify the data met the criteria for regression analysis. A scatterplot of alpha diversity, the predictor variable, against clinical efficacy, outcome variable, was constructed to provide a preliminary assessment of the relationship between the two variables. Additionally, outliers, nonlinearity, and independence of the error were evaluated and considered for compliance with the model. Data points were evaluated to assess whether they were outside the +/-3.0 standard deviation cut-off and were considered for exclusion as potential outliers.

Assumption of linearity was assessed through a plot of the residuals versus the predicted values. Scatterplots for studentized residuals of alpha diversity and probability P-P plots of the standardized residuals were produced to evaluate for constant variance. Additionally, normality

of the residuals and skewness for alpha diversity were assessed through visual evaluation of the standardized residual histogram and assessed for the need for log transformation. Fischer's measure of skewness was calculated to assess the significance if skewness was observed. Because the model for Aim 2 included only one predictor, multicollinearity was not evaluated.

Prior to conducting the analysis, the sample was assessed for missing values through frequency statistics and cross tabulation tables. Variables were assessed for inherent differences or patterns in the missing data. An application for dealing with missing data i.e., listwise and case wise removal, was considered and performed. Descriptive statistics were run to assess the impact of any data removal on the variables.

A simple linear regression was performed to test the predictor, alpha diversity, for significance with the outcome variable, efficacy rate. The regression model correlation coefficient and associated model statistics were evaluated to assess whether alpha diversity predicted the outcome. Results of the simple linear regression were evaluated through an analysis of variance (ANOVA) and standard error of the regression coefficients. Goodness of Fit was assessed through evaluation of the coefficient of determinant (R Square) to see how well the predictor data fit the model. Likelihood ratio (*F* statistic) was evaluated to assess how well alpha diversity predicts the outcome. Odds ratios were calculated and summarized for variables included in the final model.

Limitations and Threats to Validity

A benefit of conducting regression analyses utilizing retrospective data is that this approach allows the researcher to examine variables already present in the situation without having to prospectively control or manipulate the process (N. Burns & Grove, 2001). However, this approach is limited to analyses based on associations only and not causality. Additional

research would be needed in order to test for any causality between FMT donor characteristics and passing the FMT donor screening processes and 16S rRNA intestinal microbiome composition and FMT clinical efficacy of donor fecal material in treating recurrent CDI. The presence of collinearity and interactions between predictors present additional challenges to utilizing logistic regression to evaluate the effects of multiple independent predictors on the outcome, donor status. Despite these limitations, statistical steps were taken to validate associations as much as possible by minimizing alternative explanations due to bias, potential confounding factors, or random error.

As with all quasi-experimental studies, issues inherent in the sampling process may lead to selection biases. While the flexibility of these study types offers researchers options when faced with challenges related to subject recruitment, inherent differences between the groups may emerge in in substantive ways. Efforts were taken to reduce incomplete documentation and missing data by reviewing source documentation and confirming with the nursing and physician study staff. Although measures were taken to reduce sample bias, descriptive and statistical analyses were done to identify and acknowledge any inherent differences within in the study sample.

Protection of Human Subjects

Ethics board approval for the OpenBiome's research protocol was obtained from MIT on July 18, 2013. All approved study staff interacting with FMT donors received training in conducting ethical research studies and completed the online Human Subjects Research protection certification course offered through the Collaborative Institutional Training Initiative (CITI). Informed consent was obtained from prospective FMT donors who completed the FMT donor screening process and prior to donating stool for fecal transplant. During the informed

consent process, study procedures were carefully explained to each prospective FMT donor and the donor was asked to read the entire informed consent form. The study staff or investigator answered any questions presented by prospective FMT donors. Prospective FMT donors were reminded that they could choose to voluntarily withdraw from the study at any time for any reasons. Consent was reviewed again with the prospective FMT donors, prior to and throughout the intervention as needed. Participation in the study was open to adults regardless of gender and ethnicity and no harmful or hazardous procedures were conducted on the prospective FMT donors. Ethics board approval for this secondary analysis of previously collected, de-identified data was obtained on March 9, 2018 from the Boston College institutional review board for the protection of human subjects. Written documentation of permission to work with OpenBiome's study data was also obtained prior to conducting the study.

Data Protection

The research records used in this study are the property of OpenBiome. To protect participant confidentiality, the stool bank assigned prospective FMT donors with a random code upon enrollment in the study. Subsequently, all source data was referenced using the random donor number only. The key to this code linking subject identification with the random number is stored separately from the source documentation and databases. No identifying information about the FMT donors were received or used within this study. The database containing the data for this analysis were entirely de-identified and kept on a password protected computer.

Chapter 4. Results

This chapter summarizes the data collected on the subjects and results of the data analysis. The chapter is divided into two primary sections with the first addressing the analysis and results for Aim 1 and the second addressing the analysis and results for Aim 2.

Aim 1 Description of Sample and Results of Analysis

Data were collected from a total of 782 prospective donors who completed OpenBiome's prospective fecal microbiota transplantation (FMT) donor screening program between August 16, 2013 and December 30, 2016. Data collection was conducted using the methodology outlined within the chapter on study methods. Upon completion, data for the regression analyses were entered directly into an IBM® Statistical Package for the Social Sciences (SPSS) version 25 database from the Microsoft Excel© data collection tool. Once entered, the data were cleaned and reviewed for accuracy. During review of the data, 12 donors (1.5%) were found to not meet inclusion criteria and thus not eligible to participate in OpenBiome's prospective donor screening. Of the 12 excluded donors, nine donors (75.0%) were excluded based on a body mass index (BMI) of 30.0 or greater and three donors (25.0%) were excluded based on age. Data from these donors were removed from the database, bringing the final number of eligible prospective donors who met enrollment criteria and had data collected for this study to a total of 770 subjects.

Prior to conducting the analysis, the data were evaluated for any missing values. A descriptive frequency analysis demonstrated that there were a total of 54 (0.9%) missing data points affecting 26 unique prospective donors (3.4%). Results of the frequency analysis of missing data demonstrated that 15 of the prospective donors (1.9%) had no more than one missing data point, while 11 of the prospective donors (1.4%) were missing two or more data points. Cross tabulation tables were created to evaluate for inherent differences and any patterns

in the missing data. Review of the tables demonstrated that the missing values were not related to any particular variable and were missing completely at random.

Based on the analysis, the decision was made to adopt a listwise deletion approach to the missing data. This decision was supported by the observation that the missing data appeared to occur completely at random, the sample size was sufficiently large to maintain adequate power, and the number of prospective donors that contributed to the missing data represented a small portion of the total sample size (n=26, 3.4%). As such, removing these prospective donors would be unlikely to have a significant impact on the analyses and overall regression model.

Descriptive statistics were calculated for both continuous (Appendix A) and categorical variables (Appendix B) prior to and post listwise removal of the data to evaluate for any patterns. Comparison of the tables found similar results between the pre- and post-removal groups, supporting the decision to use a listwise deletion approach. Upon completion of the listwise removal of the prospective donors with missing data, the final dataset utilized for this analysis included data from 744 prospective donors.

Descriptors and Frequencies

Results from the overall prospective donor screening program were evaluated to better understand the sample population and data utilized in this study. Of the 744 prospective donors included in this study, a total of 610 (82.0%) prospective donors were deferred from donating upon completion of the prospective screening program (Table 3).

Table 3.
Overall results of OpenBiome's prospective FMT donor screening

Outcome	Total (n=744)	Clinical Assessment (n=610)	Stool Screening (n=610)	Blood Screening (n=610)
Defer	610 (82.0%)	531 (87.0%)	73 (12.0%)	6 (1.0%)
Pass	134 (18.0%)			

Of the prospective donors who were deferred, the majority of donors (n=531, 87.0%) were deferred during the clinical assessment. Donors were predominantly deferred due to a qualifying event related to mental health and well-being (n=95, 17.9%), which included a positive history for depression, anxiety, and other mental health related diagnoses (Table 4). A history of health diagnoses related to atopy, asthma, and allergies was the second most frequently reported rationale for deferral (n=92, 17.3%), followed by a history of a relevant infectious disease (n=58, 10.9%). In total, these top categories accounted for 46.1% (n=245), nearly half of the prospective donor deferrals. Further explanation of the deferral criteria utilized during the clinical assessment can be found in Appendix C.

Table 4.
Primary rationale for donor deferral during clinical assessment (n=531)

Rank	Rationale	Total n (%)
1	Mental Health and Well-Being	95 (17.9)
2	Atopy, Asthma, Allergies	92 (17.3)
3	Infectious Disease History	58 (10.9)
4	Sexual Behavior and History	53 (10.0)
5	Medications and Supplements	44 (8.3)
6	Social History	40 (7.5)
7	Gastrointestinal Conditions	30 (5.5)
8	Family History	26 (4.9)
9	Other	21 (4.0)
10	Travel History	21 (4.0)
11	Autoimmune Conditions	16 (3.0)
12	Vital Signs	14 (2.6)
13	Surgical and Other Medical History	5 (0.9)
14	Neurological Conditions	4 (0.8)
15	Cancer History	3 (0.6)
16	Cardiovascular and Metabolic Conditions	3 (0.6)
17	Diet and Exercise	2 (0.4)
18	Women's Health	2 (0.4)
19	Current Health Status	1 (0.2)
20	Musculoskeletal Conditions	1 (0.2)

During the stool and serological screening assessment, a total of 79 (13.0%) prospective donors were deferred. Health related deferrals were predominantly due to abnormal laboratory

findings (n=46, 58.2%), of which 42 (91.3%) were due to the identification of a known infectious pathogen in an otherwise asymptomatic prospective donor and 4 (8.7%) due to abnormal non-infectious serological results (Table 5). The remaining 33 prospective donors (n=33, 41.8%) were deferred during the laboratory screening phase as a result of a loss to follow-up. These prospective donors passed the clinical assessment phase but failed to complete either the blood or stool sections of the laboratory screening phase for the prospective donor screening program. The final 134 (18.0%) prospective donors who passed all steps of the prospective donor screening program were allowed to donate fecal material for use in FMT.

Table 5.
Primary rationale for donor deferral during stool and blood screening

Rank	Rationale	Total (n=79)
1	Infectious Disease	42 (53.2%)
2	Non-Infectious Serologic Test Result	4 (5.1%)
3	Lost to Follow-up	33 (41.8%)

Descriptive statistics including mean, range, standard deviation, and variance were calculated for all four continuous variables to further describe the sample population and variables of interest (Table 6). On average, prospective donors were approximately 28 years of

Table 6.
Statistical summary of continuous variables

Predictor Variable	Mean		Range	Min	Max	Std. Deviation	Variance
	Statistic	Std. Error					
Age (years)	28.1	0.23	31.0	18.0	49.0	6.19	38.25
Body Mass Index (kg/m^2)	23.6	0.10	12.9	16.9	29.8	2.69	7.22
Frequency of Bowel Movements (per day)	1.7	0.02	4.5	0.5	5.0	0.67	0.45
Alcohol Use (per week)	2.9	0.10	20.0	0.0	20.0	2.71	7.33

age, had a healthy BMI of 23.6, passed between one to two bowel movements per day (1.7 bowel movement per day on average), and consumed fewer than three alcoholic beverages per week

(2.9 beverages per week on average). In addition, data for all four categorical variables were summarized into proportions (Table 7). In general, prospective donors were found to be predominantly male (68.1%), consumed a meat-based diet (86.7%), and never smoked (81.3%). Most prospective donors were screened in the fall (29.2%) or winter (29.2%), with a notable drop in screenings observed during spring (18.6%) and moderate recovery in prospective donor screenings during summer (23.0%).

Table 7.
Statistical summary of categorical variables

Predictor Variable		Frequency	%	Cumulative %
Gender	Male	507	68.1	68.1
	Female	237	31.9	100.0
Diet	Meat Based	645	86.7	86.7
	No Meat	99	13.3	100.0
Tobacco Use	Never	605	81.3	81.3
	Former	87	11.7	93.0
	Current	52	7.0	100.0
Seasonality	Winter	217	29.2	29.2
	Spring	139	18.6	47.8
	Summer	171	23.0	70.8
	Fall	217	29.2	100.0

Statistical Analysis and Data Management

All eight variables selected *a priori* for the logistic regression in Aim 1 were evaluated to confirm model inclusion and assess for violations of the regression assumptions. Four of the variables selected *a priori* were continuous (age, BMI, frequency of bowel movements, and alcohol use) and four were categorical (gender, diet, tobacco use, and seasonality). The following section outlines the evaluation of the variables, testing of the logistic regression assumptions, and any data management steps taken.

Measures of central tendency and variance for each continuous variable were reviewed for any trends. Frequency distribution histograms and corresponding probability plots (P-P plot) for all four continuous variables were generated and evaluated for normality and skewness

(Appendix D). Data for age were predominantly normally distributed but demonstrated a slight positive skew in both the histogram and P-P plot. Data from BMI generally appeared to be normally distributed in both the histogram and P-P plot. The histograms and P-P plots for both the frequency of bowel movements and alcohol use demonstrated an obvious positive skew indicating a clustering of the data toward the low end of the scale and possible significant outliers.

Findings of skewness were further evaluated using Fisher's measure of skewness, which was calculated by dividing the skewness statistic by its standard error (Table 8).

Table 8.
Evaluation of continuous variables for skewness

Predictor Variable	Skewness Statistic	Standard Error	Fisher's Measure
Age (years)	1.113	0.090	12.367
Body Mass Index (kg/m²)	0.077	0.090	0.856
Frequency of Bowel Movements (per day)	1.130	0.090	12.556
Alcohol Use (per week)	1.380	0.090	15.333

Due to the sufficiently large sample size and previous observations from evaluation of the distribution histograms (Appendix D), a higher criterion level of p <0.001 (+/- 3.29) was utilized to assess significance in order to account for small errors inherently found in larger samples (Field, 2009; Plitchta & Kelvin, 2013). Despite the limitations observed with Fisher's measure, skewness was significant at p <0.001 for the continuous variables age, frequency of bowel movements, and alcohol use. These results were consistent with the visual observation that the variables age, frequency of bowel movements, and alcohol use contained a significant number of absolute values that were three standard deviations from the mean (+/- 3.29). As such, these variables were flagged for further evaluation of outliers and consideration for recoding or log transformation.

Scatterplots for all four continuous variables were generated against the outcome variable and evaluated for nonlinearity and need to address outliers (Appendix E). Visual assessment of the scatterplots for each continuous variable demonstrated linearity throughout and independence of the error given the binary outcome variable (pass or defer). Further, fit lines indicated that as the value of each continuous variable increased, prospective donors were more likely to be deferred. While no outliers were noted for age and BMI, several outliers were observed for frequency of bowel movements and alcohol use. As such, boxplots were generated for all four continuous variables and evaluated (Appendix F). Data for BMI were similar across the two groups and did not identify any outliers. For the continuous variables age, frequency of bowel movements, and alcohol use demonstrated several data points outside the upper quartile supporting the previous finding of skewness observed with these variables.

Notably, the boxplot for alcohol use identified one potentially problematic outlier. The original source documentation for this prospective donor was reviewed to assess any discrepancies of which there were none. Further, the impact on the regression line of removing this prospective donor from the analysis was assessed and found to be minimal. Lastly, a binary logistic regression with casewise diagnostics was run for all four variables against the dependent variable. Casewise analysis for all four continuous variables revealed that there were no outliers outside the +/- 3.0 standard deviation cut off and there did not appear to be any particular pattern to the data. Given these observations and because the data from the outlier was found to be valid, the decision was made to retain the donor in the model to preserve the integrity of the original sample population and assumptions of the model.

Based on the problematic findings of skewness associated with three of the continuous variables (age, frequency of bowel movements, and alcohol use), options were evaluated to

ensure a more equal distribution of the data across the groups. Both recoding of variables and log transformation were considered. Review of the data for the continuous variable, alcohol use, revealed that a large number of prospective donors reported zero values (n=120, 16.1%) indicating that the prospective donors did not regularly consume any alcoholic beverages each week. Because log transformation is problematic when variables contain large numbers of zero values, the decision was made to recode the variables.

The continuous variable age was distributed into four clusters based on a modified version of the United States (U.S.) Census Bureau's age distribution (United States Census Bureau, 2011). The average number of bowel movements per day (1.7 bowel movements) was rounded to a whole number (two bowel movements) and used to distribute the continuous variable frequency of bowel movements into two primary clusters in order to account for outliers and skewness observed in this variable. Lastly, the U.S. Department of Health and Human Services dietary guidelines on alcohol consumption was used as a template to code the variable alcohol use (United States Department of Health and Human Services, 2015). Based on these guidelines, alcohol use was distributed into three clusters ranging from zero consumption to moderate consumption (greater than one but less than seven drinks per week) to high consumption (seven drinks or greater per week). Differences in alcohol consumption guidelines between genders were not considered for this recoding in order to maintain homogeneity of the distribution. Because this study focused on the impact of alcohol on the intestinal microbiome and not its impact on gender differences, the decision was made to ensure alcohol exposure was consistently measured. Table 9 demonstrates how these three variables were recoded for analysis purposes.

Table 9.
Recoded continuous variables

Predictor Variable	Initial Measurement	Recoded Measurement Groups
Age	• Age (discrete) of FMT donor at the time of study participation o Mean= 28.1 o Range= 31.0 o SD= 6.19	• Recoded into four primary clusters as follows: o 1= 18 to 24 years of age (n=226, 30.4%) o 2= 25 to 29 years of age (n=272, 36.6%) o 3= 30 to 34 years of age and greater (n=146, 19.6%) o 4= 35 to 50 years of age (n=100, 13.4%)
Frequency of Bowel Movements (per day)	• The average number (discrete) of bowel movements a donor self-reports as passing each day, rounded to the first decimal. o Mean= 1.7 o Range= 4.5 o SD= 0.67	• Recoded into two primary clusters as follows: o 1= Fewer than two bowel movements per day (n=465, 62.5%) o 2= Two or more bowel movements per day (n=279, 37.5%)
Alcohol Use (per week)	• An average (continuous) of the prospective donor's self-reported weekly consumption of alcoholic beverages, rounded to the first decimal place. o Mean= 2.9 o Range= 20.0 o SD= 2.71	• Recoded into three primary clusters as follows: o 1= None; 0 drinks per week (n=120, 16.1%) o 2= Moderate; < 7 drinks per week (n=549, 73.8%) o 3= High; 7 drinks or greater per week (n=75, 10.1%)

The assumption of linearity in regression models asserts that the data points for continuous variables share a consistent and specific linear relationship throughout. Nonlinearity of the data points could result in a regression model that over- or under-estimates the value of the predictor. As such the estimator or regression must be linear. Prior to evaluating plots of the residuals, a logistic regression was run to test the assumption of linearity by evaluating for any interaction between the remaining continuous variable predictor, BMI, and its log (Table 10).

Review of the regression output demonstrated that the interactions were not significant at the $p < 0.05$ level, indicating that the assumption of linearity was met for BMI.

Table 10.
*Test for linearity of the logit**

	B	S.E.	Wald	df	*p*	Exp(B)
Body Mass Index	2.508	2.291	1.199	1	0.274	12.279
Body Mass Index by Natural Log Transformation for Body Mass Index	-0.610	0.551	1.224	1	0.269	0.544
Constant	-15.141	12.892	1.379	1	0.240	0.000

*Variable(s) entered on step 1: Body Mass Index, Body Mass Index * Natural Log Transformation for Body Mass Index.

Linearity was further assessed through visual evaluation of the residuals for the continuous variable BMI (Appendix G). Despite the limitations of evaluating scatterplots of residuals when a binary outcome variable is utilized, visual evaluation of the plots did not demonstrate any obvious abnormalities from what would otherwise be anticipated within a logistic regression model. As such, it was concluded that the plots concurred with the results observed for the regression test for linearity of the logit and supported the assumption that linearity had been met. To assess normality, a histogram of standardized residuals was generated and reviewed (Appendix H). Although one group clearly contained a higher number of data points, the data appeared to be grouped normally. Observations presented in the histogram were consistent with findings anticipated when using a binary logistic regression model. As such, normality of the data was assumed.

Chi-square testing was conducted to allow for group comparisons among the predictor variables to evaluate for differences and any relationship among the variables. Both *a priori* and recoded categorical variables were used in this analysis. The predictor variables were first examined through a correlation matrix for values greater than 0.80, which, if found, would suggest the variables could be interrelated and require further evaluation (Munro, 2000).

Correlations among several study variables were observed to be significant at the $p < 0.01$ and $p < 0.05$ (2-tailed) levels (Table 11). The variables gender and frequency of bowel movements were observed to have the highest potential correlation suggesting further inquiry was required. Similarly, alcohol use and tobacco use were noted to be significant at the $p < 0.01$ levels. Despite these findings, none of the initial correlations assessed at the 2-tailed level of significance were found to be highly correlated with correlation values greater than an absolute value of 0.80 suggesting the correlations were low (Field, 2009).

Table 11.
Correlations for predictor variables included in logistic regression

	Gender	Age	Freq of BM	Diet	Alcohol Use	Tobacco Use	Seasonality
Gender	1						
Age	0.040	1					
Frequency of Bowel Movements	-0.172**	-0.050	1				
Diet	0.055	0.008	-0.001	1			
Alcohol Use	-0.072	-0.073*	0.048	-0.094*	1		
Tobacco Use	-0.059	-0.011	0.036	-0.003	0.117**	1	
Seasonality	-0.079*	-0.038	0.062	-0.065	0.014	-0.017	1

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

Collinearity statistics were subsequently run to further evaluate the relationships among the continuous variables (Table 12). The average variance inflation factor (VIF) of the group was calculated as 1.024, which was marginally above 1.0. This indicated that multicollinearity had a low likelihood of introducing bias into the logistic regression. Similarly, this appeared to be predominantly driven by the variables, gender and frequency of bowel movements. However, evaluation of the tolerance levels and percent of variance demonstrated a low amount of variance shared by other variables, 4.3% for gender and 3.5% for frequency of bowel movements. As

such, based on the results of these analyses and taking into consideration the large sample size utilized in this study, the risk from multicollinearity was considered to be minimal.

Table 12.
Collinearity statistics

	VIF	Tolerance	
		Statistic	%
Gender	1.045	0.957	4.3
Age	1.009	0.991	0.9
Frequency of Bowel Movements	1.037	0.965	3.5
Diet	1.015	0.985	1.5
Alcohol Use	1.033	0.968	3.2
Tobacco Use	1.018	0.982	1.8
Seasonality	1.014	0.986	1.4

Aim 1 Logistic Regression

A backward elimination logistic regression was conducted to test the hypothesis that there is a statistically significant association between one or more prospective FMT donor characteristic (predictor variables) and passing the FMT donor screening program (outcome variable). Predictor variables were entered in one block and a backward stepwise elimination approach using likelihood ratio was selected. Categorical predictors were set to indicator status for coding the standard dummy variable. The first category of each categorical variable was selected as the baseline value for coding purposes. The outcome variable was coded as "0" for deferral of a prospective donor and "1" for when a prospective donor passed the screening. The alpha level of significance was set at $p < 0.05$ for the analysis. Hosmer and Lemeshow Goodness of Fit statistic, residuals, and classification plots were further selected to evaluate how well the model fit the data. Casewise listings was set to two standard deviations. Finally, a confidence interval of 95% was used to calculate the odds ratios for significant predictors. Results of the logistic regression analysis are provided in Appendix I. All selected cases were utilized in the analysis and there were no missing cases.

The log-likelihood value of the baseline model where only the constant was included was 701.68 (Block 0: Beginning Block), which represented the most basic fit of the model to the data. For the backward stepwise elimination model with likelihood ratio, a total of eight iterations were performed using a 0.001 threshold as the parameter estimate before arriving at the final model. The eighth and final iteration of the backward elimination model included the variable frequency of bowel movements of two or more per day and reported a log-likelihood value of 695.66. The log-likelihood for the final model was reduced in comparison to the baseline model, suggesting that when donor frequency of bowel movements was included, the model was better at predicting whether a prospective donor would pass or be deferred. Analysis of the model chi-square statistic of 6.017 taken from the Omnibus tests of the model coefficients was significant with a p value of 0.014 (p >0.05) providing further support for the that the results of the regression model was significantly different from the baseline model.

Notably, the classification table for the final model remained unchanged from the baseline model, with the model chosen that predicted donor outcome based on the majority of observations. Based on this approach, the model predicted that prospective donors who had a frequency of bowel movements of two or more per day would be more likely to be deferred. A cross tabulation table was generated for the predictor variable frequency of bowel movement and the outcome variable FMT donor status, defined as defer or pass upon completion of the screening program (Table 13). A total of 279 prospective donors with a frequency of bowel movements of two or more per day were deferred. Of these 279 donors, the model correctly predicted that 241 (86.4%) of these donors were deferred and misclassified 38 (13.6%) of the prospective donors.

Table 13.
Cross tabulation of FMT donor status and frequency of bowel movements

		Frequency of Bowel Movements		Total
		Fewer than two	Two or more	
FMT Donor Status	Defer	369 (79.4%)	241 (86.4%)	610 (82.0%)
	Pass	96 (20.6%)	38 (13.6%)	134 (18.0%)
	Total	465	279	744

The value of the beta coefficient (B) for the final model was -0.501 (Table 14), which had an associated Wald statistic of 5.73 and was significant at a *p* value of 0.016 ($p > 0.05$). As such, inclusion of the predictor frequency of bowel movements in the model was a significant predictor of whether a donor was deferred upon completion of the prospective donor screening program. The negative beta coefficient indicated that prospective donors with a frequency of bowel movements of two or more per day were less likely to pass the screening than donors with a frequency of bowel movements of fewer than two per day. Review of the final model log likelihood statistic reported a change of 6.017, which was significant with a *p* value of 0.014 ($p < 0.05$). As such, removal of the variable frequency of bowel movements of two or more per day would have a significant effect on the predictability of the model. The residual chi-square overall statistic for the final model was 11.703 but was not significant with a *p* value of 0.552 ($p > 0.05$) supporting the conclusion that addition of the other variables would not increase the model's predictive power. This was further supported by absence of any significant Roa's efficient score statistic values provided by the table for variables not included in the model.

Table 14.
Results of final logistic regression analysis

Variable	Beta	S.E.	Wald	*p*	Exp(B)	95% CI for Exp(B)	
						Lower	Upper
Frequency of BM							
Two or more (per day)	-0.501	0.209	5.753	0.016	0.606	0.403	0.913
Constant	-0.846						

The odds ratio, labeled Exp (B), is an indicator for the change in odds as a result of a one-unit change in the predictor. For the final model predictor, frequency of bowel movements of two or more per day, the odds ratio was 0.606 with an associated confidence interval of 0.403 to 0.913 (Table 14). Given the value of 0.606 change in odds was less than one, the model suggested that the odds of the outcome occurring, decreased with an increase in the value of the predictor. Namely, the odds of a donor passing the prospective donor screening with a frequency in bowel movements of two or more per day was lower at 0.606. Probabilities were also calculated to evaluate the odds for each outcome (Table 15). In summary, the odds of a donor having a frequency of bowel movements of two or more per day and passing the prospective donor screening were 13.6%. These odds were lower as compared to the 20.6% odds for donors who reported a frequency of bowel movements of fewer than two per day. The predicted probabilities concurred with the probabilities supplied in the Case summaries table resulting from the final model.

Table 15.
Final model probabilities

	Frequency of Bowel Movements	
	Fewer than two (%)	**Two or more (%)**
Pass	20.6	13.6
Defer	79.4	86.4

Evaluation of Cook's Distance (range 0.001 to 0.023) for the model found that all the variables were less than one, indicating that no specific case had an undue influence on the model. Similar findings were observed by evaluation of the Centered Leverage value (range 0.002 to 0.004) providing further evidence that there were no influential cases within the model. DFBeta statistics for the model were evaluated to assess whether any cases had a large influence on the regression parameters. The review found no absolute values greater than one for either DFBeta statistics for the constant (range 0.000 to 0.010) or frequency of bowel movements

(range 0.003 to 0.026), suggesting that no cases exerted undue influence on the regression parameters utilized in this model. Finally, a forced entry logistic regression model was rerun with the predictor variable, frequency of bowel movements, to remove any influence that the non-significant predictors may have had on the model. Results of the forced entry logistic regression concurred with the findings from the backwards elimination logistic regression analysis. Similarly, forced entry simple logistic regression models were performed for each non-significant variable (Appendix J). Results of these logistic regression analyses concurred with the findings from the backward elimination logistic regression analysis.

For the Aim 1 logistic regression, all predictor variables were entered into the model and evaluated stepwise to see if they contributed to the regression equation (Table 16). Predictor variables found not to contribute statistically to the model (reporting a *p* value equal to or greater than 0.05) were dropped until only statistically significant predictors remained. The final model contained the predictor, frequency of bowel movements two or more per day, and accounted for 0.8% (0.008 Cox & Snell R Square) to 1.3% (0.013 Nagelkerke R Square) of the variance.

Table 16.
Results of the logistic regression analysis

Variables (Step 1)	Beta	S.E.	Wald	*p*	Odds Ratio	CI 95%	
						Lower	Upper
Gender							
Female	-0.341	0.224	2.311	0.128	0.711	0.459	1.104
Age							
18 to 24 years of age			1.477	0.688			
25 to 29 years of age	-0.251	0.24	1.095	0.295	0.778	0.487	1.245
30 to 34 years of age	-0.115	0.276	0.173	0.678	0.892	0.519	1.533
35 to 50 years of age	-0.322	0.341	0.895	0.344	0.724	0.371	1.413
Body Mass Index	-0.028	0.039	0.509	0.475	0.973	0.902	1.049
Frequency of BM							
Two or more per day	-0.509	0.215	5.592	0.018*	0.601	0.394	0.917
Diet							
No Meat	-0.461	0.322	2.054	0.152	0.631	0.336	1.185

Table 16. (continued)

Variables (Step 1)	Beta	S.E.	Wald	*p*	Odds Ratio	CI 95%	
Alcohol Use							
None			0.459	0.795			
Moderate	-0.027	0.267	0.01	0.921	0.974	0.577	1.645
High	-0.249	0.405	0.378	0.539	0.78	0.353	1.724
Tobacco Use							
None			1.121	0.571			
Former	-0.11	0.319	0.119	0.730	0.896	0.480	1.673
Current	-0.438	0.428	1.048	0.306	0.645	0.279	1.493
Seasonality							
Winter			4.244	0.236			
Spring	-0.491	0.309	2.528	0.112	0.612	0.334	1.121
Summer	-0.297	0.277	1.148	0.284	0.743	0.432	1.279
Fall	0.045	0.243	0.035	0.852	1.047	0.650	1.685
Constant	0.341	0.997	0.117	0.732	1.406		

* Significant at the *p* value level of <0.05

Overall, the final model was found to be more predictive of donor passage or deferral when the variable frequency of bowel movements was included. However, despite the statistically significant finding, further evaluation of the model parameters supported that the overall model was not a good fit or very good at predicting the outcome. Evaluation of Hosmer and Lemeshow Goodness of Fit test was not feasible due to inherent constraints resulting from the use of a binary predictor in the final model (degrees of freedom equal to 0). In summary, frequency of bowel movements, while found to be a significant predictor, did not explain most of the variability in the outcome. Most notably, the remaining seven predictors including gender, age, BMI, diet, alcohol use, tobacco use, and seasonality were not statistically significant predictors of donor passage or deferral based on the logistic regression despite the theoretical underpinnings suggesting that these variables may play a role in overall donor health.

Aim 2 Description of Sample and Results of Analysis

Data were collected from a subset of screened prospective donors for the analysis associated with the second aim of this study. The sample included prospective donors who had 1) passed the FMT donor screening program; 2) submitted a fecal sample for 16S ribosomal ribonucleic acid (rRNA) sequencing; and 3) donated material that was used in a FMT to treat patients with recurrent *Clostridium difficile* infections (CDI). Although a total of 134 prospective donors passed the screening, only 87 donors (64.9%) contributed fecal material for which clinical efficacy data were available. The clinical efficacy rate for each of the sample donors was entered into a Microsoft Excel© spreadsheet as a cumulative percentage rate (rounded to the nearest tenth), verified for accuracy, and subsequently uploaded into an IBM® SPSS version 25 database for analysis.

One copy of 16S rRNA sequencing data for each of the sample donors was requested and retrieved from the data repository. Fecal microbiome 16S rRNA sequencing data were available for all 87 of the requested donors. A compressed file containing de-multiplexed paired strands was downloaded onto an Apple Macintosh® computer capable of processing the raw sequencing data using QIIME 2© version 2018.8 microbiome data analysis software (Caporaso et al., 2011; QIIME 2 development team, 2018). De-multiplexed paired strands for all sample donors were imported into the QIIME 2 software without error. A summary table was generated to visualize the data for sequencing quality control (Appendix K). Based on visual inspection, the quality of the sequence data observed was high and no data trimming was required at the beginning or end of the sequences. As such, a right trim point at 200 bases was selected based on an observed decrease in quality reads past this point. This information was used to run the Divisive Amplicon Denoising Algorithm 2 (DADA2) pipeline to detect and filter out biological variations in the

sequencing data associated with amplicon sequencing errors, including phiX reads and chimeric sequences (Callahan et al., 2016). The results of this quality filtering were used to construct the Feature Table and Feature Data files needed to evaluate the donor fecal microbiome data. No errors or issues were observed in the sequencing data upon completion of the quality control measures executed in QIIME 2.

Descriptors and Frequencies

Descriptive statistics were calculated using the same variables utilized during Aim 1 to better understand and evaluate the sample population used in the analysis for Aim 2 (Table 17). On average, the donors included in the sample for Aim 2 were approximately 27 years of age, had a healthy BMI of 23.2 kg/m^2, passed fewer than 2 bowel movements per day (1.6 bowel movements per day on average), and consumed approximately 3 alcoholic beverages per week (3.1 beverages per week on average).

Table 17.
Statistical summary of continuous variables

Predictor Variable	Mean		Range	Min	Max	Std. Deviation	Variance
	Statistic	Std. Error					
Age (years)	27.3	0.6	26.0	19.0	45.0	5.48	30.00
Body Mass Index (kg/m²)	23.2	0.3	11.6	17.5	29.1	2.55	6.49
Frequency of Bowel Movements (per day)	1.6	0.1	3.0	0.5	3.5	0.67	0.45
Alcohol Use (per week)	3.1	0.3	10.0	0.0	10.0	2.47	6.12

Data for all four categorical variables were summarized into proportions and evaluated for trends (Table 18). In general, the sample population used in Aim 2 was predominantly male (72.4%), consumed a meat-based diet (90.8%), and never smoked (83.9%).

Table 18.
Statistical summary of categorical variables

Predictor Variable		Frequency	%	Cumulative %
Gender	Male	63	72.4	72.4
	Female	24	27.6	100.0
Diet	Meat Based	79	90.8	90.8
	No Meat	8	9.2	100.0
Tobacco Use	Never	73	83.9	83.9
	Former	8	9.2	93.1
	Current	73	83.9	100.0
Seasonality	Winter	29	33.3	33.3
	Spring	11	12.6	46.0
	Summer	17	19.5	65.5
	Fall	30	34.5	100.0

Descriptive statistics for both the continuous variables (Appendix L) and categorical variables (Appendix M) utilized in the sample population for Aim 2 were compared with the sample population utilized in Aim 1 to evaluate for any obvious patterns or discrepancies. Despite the notably smaller sample size in Aim 2, comparison of the tables found similar results between the two sample populations used in Aim 1 and Aim 2.

Microbiome Analysis

A metadata file was generated and imported to QIIME 2, which included the identifier, sequence information, and cumulative clinical efficacy for each donor. The variable found to be significant from Aim 1, frequency of bowel movements, was included in the metadata file. The information contained within the metadata, Feature Table, and Feature Data files were used to explore the fecal microbiome composition and characteristics for the total sample and by individual donor fecal material. This information was further used to evaluate the beta and alpha diversity of the total sample and by individual donor.

Characteristics of the fecal microbiome for the overall sample. Descriptive statistics were calculated at the bacterial phylum level to evaluate the frequency of features e.g., bacterial operational taxonomic units (OTU), present within the sample (Table 19). Bacterial phyla were

organized in descending order from most to least frequent. The total frequency of bacterial OTUs observed for the entire sample, all 87 donors, was 2,769,977 with a per donor frequency rate ranging from a low of 5,705 OTUs to a high of 71,173 OTUs. The average frequency of OTU bacterial features for the sample was 31,839 for the sample. Within the total sample population, the most abundant bacteria represented belonged to the *Firmicutes* phylum (n=1,632,950, 59.0%) followed by bacteria from the *Bacteroidetes* phylum (n=897,287, 32.4%). Overall, bacteria from the *Firmicutes* and *Bacteroidetes* phyla contributed 91.3% of the frequency of bacterial OTU features observed within the sample.

Table 19.
Statistical summary of sample frequencies or OTUs by phylum

Phylum	N	%	Mean	Range	Min	Max
Firmicutes	1,632,950	59.0	18,770	41,865	3,990	45,855
Bacteroidetes	897,287	32.4	10,314	43,477	1,134	44,611
Actinobacteria	127,813	4.6	1,469	9,933	62	9,995
Proteobacteria	68,874	2.5	792	10,532	74	10,606
Verrucomicrobia	19,481	0.7	224	2,969	0	2,969
Euryarchaeota	11,230	0.4	129	1,558	0	1,558
Cyanobacteria	5,062	0.2	58	715	0	715
Fusobacteria	4,746	0.2	55	2,768	0	2,768
Tenericutes	1,860	0.1	21	253	0	253
Lentisphaerae	393	0.0	5	65	0	65
Bacteria, unclassified	220	0.0	3	198	0	198
Synergistetes	58	0.0	1	19	0	19
Spirochaetes	3	0.0	0	3	0	3
Total	2,769,977	100.0	31,839	65,468	5,705	71,173

The next nearest phylum, *Actinobacteria*, was notably lower contributing only 4.6% (n=127,813), while the *Proteobacteria* phylum contributed an additional 2.5% of the bacterial OTU features. In total, the four aforementioned bacterial phyla, *Firmicutes*, *Bacteroidetes*, *Actinobacteria*, and *Proteobacteria* contributed 98.4% of the feature OTUs found in the overall sample.

Characteristics of the fecal microbiome by individual donor. Descriptive statistics were generated at the individual donor level to evaluate and better understand the fecal

microbiome composition contributed by each donor and to observe for any differences. A histogram presenting the frequency in number of bacterial OTU features provided by each donor sample demonstrated a predominantly normal shaped distribution across the sample (Appendix N). A summary table containing descriptive statistics at the phylum level was generated using these data to visually evaluate community richness by donor (Table 20). Consistent with the overall observations for the sample, the most prominent phylum, *Firmicutes*, ranged between 22.3% to 83.6% of the composition of the fecal microbiome of each donor and, on average, composed 59.2% of a donor's fecal microbiome. The second most abundant bacterial phylum, *Bacteroidetes*, ranged in composition from 10.9% to 62.7% of the fecal microbiome for each donor and, on average, comprised 32.0% of a donor's fecal microbiome.

Table 20.
Summary of percentage composition of bacterial phylum by donors

Predictor Variable	Mean		Range	Min	Max	Std. Deviation	Variance
	Statistic	Std. Error					
Firmicutes	59.2	1.015	61.3	22.3	83.6	9.470	89.689
Bacteroidetes	32.0	0.979	51.8	10.9	62.7	9.131	83.380
Actinobacteria	4.6	0.346	14.8	0.2	15.0	3.230	10.435
Proteobacteria	2.5	0.287	23.8	0.3	24.1	2.680	7.184
Verrucomicrobia	0.8	0.170	8.8	0.0	8.8	1.586	2.516
Euryarchaeota	0.4	0.102	3.9	0.0	3.9	0.950	0.902
Cyanobacteria	0.2	0.063	3.1	0.0	3.1	0.588	0.346
Fusobacteria	0.2	0.127	10.5	0.0	10.5	1.188	1.412
Tenericutes	0.1	0.021	1.0	0.0	1.0	0.200	0.040
Lentisphaerae	0.0	0.004	0.2	0.0	0.2	0.038	0.001
Bacteria, unclassified	0.0	0.006	0.5	0.0	0.5	0.054	0.003
Synergistetes	0.0	0.000	0.0	0.0	0.0	0.000	0.000
Spirochaetes	0.0	0.000	0.0	0.0	0.0	0.000	0.000

An OTU taxonomy bar chart was generated to visually evaluate the per donor observations summarized in the table (Appendix O). Consistent with the previous findings, bacteria from the phyla *Firmicutes* and *Bacteroidetes* were predominantly observed in the compositions for each donor. On the low end of the scale, bacteria from the combined groups,

Firmicutes and *Bacteroidetes*, composed 72.2% of the donor's fecal microbiome. For this donor, bacteria from the *Proteobacteria* phylum contributed 24.1% of the remaining fecal microbiome. In contrast, the fecal microbiome of the donor with the second lowest combined percentage composition of *Firmicutes* and *Bacteroidetes* (80.6%) was comprised primarily of bacteria from the *Actinobacteria* phylum, which composed 15.0% of the donor's fecal microbiome. On the high end of the table, bacteria from the combined group *Firmicutes* and *Bacteroidetes* comprised 98.2% of the donor's fecal microbiome. One donor had a notably higher percentage composition of bacteria from a less common phylum, *Fusobacteria* (10.5%). Bacteria from this phylum were otherwise not well represented elsewhere in the sample population.

A heatmap based on donor 16S rRNA data can be useful to explore for any unusual patterns in donor contribution and corroborate the observations with those presented by the OTU taxonomy bar chart. As such, a heatmap was generated at the phylum level to further evaluate the donor sample (Appendix P). Generally, most donors were observed to contribute similar OTU frequencies across all phyla. However, bacteria from the phyla *Verrucomicrobia* and *Euryarchaeota* comprised higher portions of the intestinal composition for a select subset of donors. When compared with the taxonomy bar chart, this same observation was also noted in OTUs (Appendix O). On further exploration, these two phyla groups combined (*Verrucomicrobia* and *Euryarchaeota*) contributed less than 10% of the overall frequency of OTUs for any of the donors.

A Principal Coordinate Analysis plot (PCoA) was generated using the QIIME 2 software platform to visually evaluate the microbiome diversity between donors, referred to as beta diversity. Beta diversity is utilized to evaluate how different two or more microbiome communities are from each other and represents a between samples analysis. Using a Bray-Curtis

approach and taxonomic data provided by the16S rRNA donor data, relative distance matrices for donors included in the sample were calculated using the QIIME 2 software platform and plotted on a three-dimensional emperor plot (Appendix Q). Each data point in the PCoA plot represented an individual donor. The resulting PCoA plot demonstrated an overall clustering of the donor samples within a large data cloud, indicating that most donors shared similar microbiome habitats. Variance explained by the three axes in the PCoA plot ranged from 5.6% to 11.3%. Despite the low values, overall observation suggested that donors included in the sample did not differ significantly (Goodrich et al., 2014; Kuczynski et al., 2010). Two donors in particular, however, appeared to be notably different from the remainder of the donors. Despite a review of these donors' data, there were no apparent dissimilarities in the donor characteristics, nor were any data observed that could account for the differences observed in microbial habitat.

The Shannon diversity index was utilized to evaluate how diverse each donor sample was independently, referred to as alpha diversity. Alpha diversity as measured by the Shannon diversity index provides a measure of both the abundance and evenness of the taxa that are present within a specific donor sample (Shannon, 1997). This approach demonstrates a within sample evaluation and tends to vary across different sampled sites. In general, fecal samples tend to be more diverse and present higher alpha diversity levels than other sampled human microbiome sites (J. Peterson et al., 2009). The QIIME 2 software platform provides several measures of alpha diversity. However, for this study the data represented by the Shannon diversity index were selected for further evaluation (Table 21). On average, the alpha diversity of the donor sample was 5.5 and ranged from a low of 2.6 to a maximum of 6.4. No apparent variations were noted in the data results and overall observations were consistent with donor data

reported in the literature on FMT donors (Kelly et al., 2016; Seekatz et al., 2014; van Nood, Dijkgraaf, et al., 2013).

Table 21.
Statistical summary of continuous variable, alpha diversity

Predictor Variable	Mean		Range	Min	Max	Std. Deviation	Variance
	Statistic	Std. Error					
Alpha Diversity	5.5	0.055	2.6	3.8	6.4	0.509	0.259

Shannon's diversity index for each donor was used to generate a box and whisker plot to evaluate for any difference in alpha diversity among the categorical variable, frequency of bowel movements (Appendix R). This variable was found to be predictive of a donor passing the FMT donor screening program. Visual observation demonstrated that donors who passed fewer than two bowel movements per day did appear to have a higher level of alpha diversity as compared with donors who had two or more bowel movements per day. Descriptive statistics were calculated to further explore for any differences between the groups (Table 22). More than half of all donors who passed fewer than two bowel movements per day had an alpha diversity greater than 5.6 as compare to the 5.3 mean alpha diversity for donors who passed two or more bowel movements per day. The difference in mean alpha diversity between groups, however, was marginal and concurred with observations demonstrated by Vandeputte and colleagues' (2016) assessment of alpha diversity based on stool consistency using the Bristol stool scale.

Table 22.
Statistical summary of alpha diversity by donor frequency of bowel movements per day

	n	Mean		Range	Min	Max	Std. Deviation	Variance
		Statistic	Std. Error					
Fewer than two BM per day	63	5.6	0.052	2.6	3.8	6.4	0.416	0.173
Two or more BM per day	24	5.3	0.133	2.4	3.9	6.3	0.649	0.422

One outlier was noted in the group of donors who passed fewer than two bowel movements per day. This donor was removed from the subgroup of donors and descriptive

statistics were repeated to evaluate any changes in the data (Table 23). After excluding the potential outlier, donors who passed two or more bowel movements per day demonstrated a much wider degree of alpha diversity (range of 2.6) ranging from a minimum of 3.9 to a high of 6.3 as compared to a range of 1.6 for donors who passed fewer than 2 bowel movements per day. Despite the limited impact that removing this donor had on a change in the mean alpha diversity (5.6 pre-removal versus 5.7 post-removal), this donor was flagged for further evaluation during the linear regression analysis portion of this study.

Table 23.
Revised statistical summary of alpha diversity by donors who passed fewer than two bowel movements per day

	n	Mean		Range	Min	Max	Std. Deviation	Variance
		Statistic	Std. Error					
Fewer than two BM per day	62	5.7	0.043	1.6	4.8	6.4	0.342	0.177

Statistical Analysis and Data Management

A simple linear regression approach was chosen for Aim 2 to evaluate the relationship between the *a priori* selected continuous predictor variable, alpha diversity, and the donor's FMT clinical efficacy rate recorded as an overall percentage. Prior to the analysis, the regression variables, alpha diversity and efficacy rate, were plotted in a simple scatterplot to assess the general relationship between the two variables (Appendix S). Based on the scatterplot, alpha diversity appeared to be positively correlated with the rate of efficacy, suggesting that as alpha diversity of the donor fecal microbiome increased, the efficacy rate of the fecal material in preventing *Clostridium difficile* recurrence also increased. The two aforementioned potential outliers observed during the visual evaluation of the sample and donor microbiome were also evaluated against the scatterplot to verify if the trend continued when the efficacy rate was included. Despite the findings observed in the descriptive analysis of the donor microbiome, the

two potential donor outliers were found to reside within the large data cloud, suggesting that for the linear regression analysis, these donors were not significant outliers.

The independent variable, alpha diversity, was further evaluated to confirm the choice of the model and assess for any violations of the regression assumptions. A frequency histogram and P-P plot were generated to assess the predictor variable, alpha diversity, and evaluate measures of central tendency (Appendix T). Both figures demonstrated an obvious negative skew to the data with a large clustering of the data at the upper end of the scale. Fisher's measure was subsequently run to further evaluate the observation (Table 24). Despite the limitations of Fisher's measure, the results were consistent with a negative skewing of the data. The measure was noted to be significant at $p < 0.001$ (+/- 3.29). Based on these results, the decision was made to perform a reverse score log transformation on the predictor, alpha diversity, to address the negative skewing.

Table 24.
Evaluation of continuous variable for skewness

Predictor Variable	Skewness Statistic	Standard Error	Fisher's Measure
Alpha Diversity	-1.226	0.258	-4.752

A frequency histogram and P-P plot were rerun utilizing the reverse score log-transformed predictor variable, renamed alpha diversity log-transformed, to evaluate the impact of the reverse score log transformation on the variable's skewness (Appendix U). Both figures demonstrated a notable decrease in the level of skewness observed. Evaluation of the histogram for the log-transformed alpha diversity variable was consistent with trends that would be anticipated when utilizing a reverse score approach, namely the reversal of scores, where big scores become smaller and vice versa (Field, 2009). Although the histogram of the log-transformed variable continued to demonstrate some skewing, the skew was less prominent than

what had been observed in the previous, non-transformed histogram. Visual evaluation of the P-P plot provided similar findings with a continued, yet notably reduced skewing in the data. Fisher's measure was rerun to further evaluate the findings (Table 25). Results of the analysis demonstrated that despite the slight positive skewing, the observations were no longer significant at the p <0.001 (+/- 3.29).

Table 25.
Evaluation of transformed continuous variable for skewness

Predictor Variable	Skewness Statistic	Standard Error	Fisher's Measure
Alpha Diversity, log-transformed	0.314	0.258	1.217

A post-transformation scatterplot of the predictor variable, alpha diversity, against the outcome variable, efficacy rate, was constructed to reassess the relationship between the variables, identify any new outliers, and assess for nonlinearity (Appendix V). Visual assessment of the scatterplots demonstrated linearity throughout and supported independence of the error. Further, the scatterplot continued to demonstrate a positive relationship between alpha diversity and efficacy rate. As such, the same relationship was observed between the pre- and post-transformed scatterplots of the predictor variable alpha diversity and efficacy rate. Otherwise, there did not appear to be any particular pattern to the data cloud present in the post-transformed scatterplot indicating that a violation of the assumption of linearity was unlikely. Confidence intervals at 95% were added to the post-transformed scatterplot to assess for potential outliers that may exert an undue influence on the regression analysis (Appendix W).

Several outliers not previously flagged during the analysis were observed and marked for further evaluation. Casewise diagnostics were run and confirmed the presence of one donor residing outside the +/- 3.0 standard deviation cut off of particular concern (Appendix X). This donor had a high alpha diversity score with corresponding low clinical efficacy rate. Source

documents were requested and queried for this donor and were found to be valid. The impact to the dataset and regression line of removing this suspected donor outlier was assessed by plotting the pre- and post-removal regression lines on a scatterplot (Appendix Y). Although a shift in the regression line was observed, the change was noted to be small. Based on these findings, the decision was made to retain the donor's data within the dataset in order to maintain the integrity of the data and assumptions of the regression model.

Plots of the residuals with the predicted values for the independent variable alpha diversity log-transformed was generated to further evaluate for violations in the regression assumption (Appendix Z). The assumption of linearity was assessed through a plot of the residuals versus the predicted values. Although the graph demonstrated a dense data cloud and one suspected data point outside +/-3.0 standard deviation, there did not appear to be a particular pattern to the data indicating that a violation of the assumption of linearity was unlikely. Scatterplots demonstrated data that were generally well spread throughout the plots. A P-P plot of the standardized residuals demonstrated some bowing out of the data particularly in the middle and towards the upper end, closer to 1.0. However, the bowing did not appear to be significant, indicating that while some level of violation of constant variance might have been present, this was not likely to notably impact the regression analysis. Analysis of the residual scatterplots for the independent variable, alpha diversity log-transformed, did not demonstrate any particular pattern to the data. As such, it was assumed that independence of the error was not violated. To assess normality, a histogram of standardized residuals was generated and reviewed (Appendix AA). Although the data were predominantly normally distributed, the histogram did demonstrate some skewing to the left with a number of variables lying beyond the -4.0 point. However, this appeared to be a small portion of the overall residual.

Aim 2 Linear Regression

A simple linear regression assessing the relationship between the clinical efficacy rate and alpha diversity of the donor fecal material as measured by Shannon's diversity was used to evaluate if there was any level of predictability between the two variables. Because only one predictor variable was considered, alpha diversity, the model was set to enter all variables at once. The results of the regression analysis were calculated in model 1 as part of the test for interaction (Appendix BB). For this analysis, the alpha level of significance was set at $p < 0.05$. Residuals and classification plots were further selected to evaluate how well the model fit the data. Casewise listing was set to two standard deviations. Lastly, a confidence interval of 95% was set in order to calculate the odds ratios if the predictor, alpha diversity, was found to be significant. All cases in the sample were utilized in the regression analysis and there were no missing cases from the analysis.

On average, the efficacy rate for the overall sample of 87 donors was 78.9%, suggesting a CDI recurrence rate of 21.1%. Evaluation of the reported correlation value (R) of 0.159 observed between the two variables used in the simple regression model, alpha diversity and efficacy (Table 26), was not significant with a p value of 0.07 ($p > 0.05$). The proportion of variance (R^2) explained by the simple linear regression model was reported as 0.025, indicating that approximately 2.5% of the variability in the efficacy rate could be explained by changes in the alpha diversity of the donor's fecal microbiome. In total, 97.5% of the variability in the efficacy rate could not be explained by the predictor, alpha diversity, indicating that other variables not included in the model were likely more influential of the outcome. The adjusted R^2 value provided a more conservative estimate of 0.014 or 1.4% of explained variance in the efficacy rate. However, this was not notably different from the R^2 of 2.4% indicating that the R^2 value

was generally reflective of the model. Because this model only utilized one predictor variable, this similarity between the statistics was not unexpected.

Table 26.
Model summary[b]

Model	R	R Square	Adj R Square	Std. Error of the Estimate	R Square Change	*F* Change	Sig. *F* Change
1	0.159[a]	0.025	0.014	11.405	0.025	2.217	0.140

a Predictors: (Constant), Alpha Diversity log transformed
b Dependent Variable: Efficacy %

The standard error of the estimate associated with the reported R^2 value in the model was 11.405 and provided an estimate of the accuracy of the predictions made by the regression line. Approximately 68.0% of the scores in the model were 11.67 points above and below the predicted value. This wide spread was not surprising given the number of observations that were not explained by the model (97.5%). The observed *F* change statistic for the R^2 value in the model was reported as 2.217 with an associated level of significance *p* value of 0.14 ($p > 0.05$). Based on these results, it was concluded that the variable alpha diversity was not a significant predictor of the efficacy rate of FMT.

Analysis of the overall fit of the model was presented in the analysis of variance (ANOVA) table. The regression model's Sum of Squares statistic was reported as 288.3 and had an *F*-ratio statistic of 2.22. This statistic was found to not be significant with a *p* value of 0.14 ($p > 0.05$). Based on this result, the ANOVA table supported the conclusion that the model, efficacy rate regressed on the predictor alpha diversity, was not significantly correlated. In summary, the model found that the degree of alpha diversity of the donor's fecal material was not a predictor of the efficacy of the FMT procedure in preventing CDI recurrence.

Per the coefficients table, the regression beta coefficient value (B) obtained for the simple regression model was -16.3. This suggested that for a unit change in alpha diversity, the efficacy

rate increased by 16.3%, keeping in consideration the reversal of values resulting from the reverse log transformation of the alpha diversity data points. However, the *t* observed value of -1.49 for the regression coefficient had an associated level of significance *p* value of 0.14, which was not significant ($p > 0.05$). As such, this finding suggested that the regression coefficient value likely occurred by chance. Odds ratios were not generated because the results of the simple linear regression were found not to be significant.

Cook's Distance for the data used in the model were observed to all be less than one, ranging from 0.000 to 0.165, indicating that no specific case had an undue influence on the model. Similar findings were observed by the Centered Leverage value (range 0.000 to 0.088). DFBeta statistics for the model was predominantly less than one, though three points were noted to be slightly above the cut-off value (range 0.001 to 1.222). Evaluation of these cases demonstrated corresponding Cook's Distance and Centered Leverage values that were less than 0.165, suggesting that these cases were not likely to have significantly influenced the results of the regression analysis. A simple linear regression was run with these cases excluded to further evaluate the findings. This analysis demonstrated similar results as those observed in the final model (model 1). As such, it was concluded that no cases exerted a substantial or undue influence on the regression parameters utilized in the final model (model 1).

For the Aim 2 simple linear regression, the predictor variable alpha diversity was entered into the model and evaluated to see if it contributed to the regression equation. The final model (model 1) accounted for only approximately 2.5% of the variance in the outcome variable. Efficacy rate and was found to not be significant with a *p* value of 0.14 ($p > 0.05$) (Table 27). As such, it was determined that the level of alpha diversity of a donor's fecal microbiome was not predictive as to whether the material was efficacious in preventing the recurrence of CDI.

Table 27.
Results of the simple linear regression analysis[a]

Model	Beta	S.E.	Std. Coefficients Beta	*t*	*p*	CI 95%	
						Lower	Upper
(Constant)	83.095	3.050		27.245	0.000	77.031	89.159
Alpha Diversity	-16.337	10.973	-0.159	-1.489	0.140	-38.154	5.481

a Dependent Variable: Efficacy %

Overall, evaluation of the donor microbiome for Aim 2 demonstrated a predominantly homogenous sample comprised of bacteria from the *Bacteroidetes* and *Firmicutes* phylum (70% or greater). Despite indications of some variability observed in the beta analysis, the simple regression analysis was not significant, indicating that the observed variability did not impact the clinical efficacy of the donor material. Additionally, the non-significant result of the regression analysis suggested that any observed relationship between donor alpha diversity and clinical efficacy of the material might have occurred by chance.

Chapter 5. Discussion

There is a paucity of research on ideal healthy stool donor characteristics and the microbiome of donor fecal material. Because of the lack of consensus-based data to support a correlation between donor characteristics and passing the prospective donor screening, nurses and other clinicians face multiple hurdles and operational challenges in finding stool donors to treat recurrent *Clostridium difficile* infections (CDI) patients with consistently efficacious fecal microbiota transplant (FMT) material. This research set out to build knowledge about an emerging area of science and clinical care. Two hypotheses aimed at improving the process of finding suitable donors with consistently efficacious intestinal microbiomes were tested. The first aim evaluated eight variables for any correlation with passing OpenBiome's prospective FMT donor screening program. The second aim evaluated a subset of donors who passed the screening to determine any correlation between the stool donor's intestinal microbiome composition (as measured by 16S ribosomal ribonucleic acid [rRNA] based Shannon's alpha diversity index) and the clinical efficacy of FMT. The results, along with study limitations, implications for clinical practice, and recommendations for future research are discussed in this chapter.

Sample Representation of Stool Donor Population

Prior to conducting the analyses, the characteristics of the sample were compared with information about stool donors reported in the literature. Overall, the study sample was similar to samples described in the literature. However, in this study, 18.0% (n=134) of the donors passed OpenBiome's prospective FMT donor screening program, a rate that was notably higher than those reported for other screening programs, which ranged from 4% (L. J. Burns et al., 2015) to 10% (Paramsothy et al., 2015). This difference may be explained by the fact that donors in this study had already been prescreened using a targeted online questionnaire, whereas in other

studies, stool donors were recruited directly from the general population. The online tool used by OpenBiome had been tailored to capture common deferral criteria prior to scheduling an in-person assessment, such as a high body mass index (BMI), recent use of antibiotics, and age outside the inclusion range. As such, the prescreening may have prospectively deferred a number of donors prior to the more rigorous in-person clinical and laboratory screenings and account for the difference in passage rates observed in the literature. Despite this difference, the sample cohort utilized in this study did not appear to be notably different from those reported in the literature (L. J. Burns et al., 2015; Paramsothy et al., 2015) and thus was considered to be a representative sample of the pool of stool donors. Although the online prescreening was not evaluated as part of this study, these results suggested that the use of a short online survey utilizing targeted questions might be a quick and economic approach to improving the FMT prospective donor pass rates.

Discussion of the Findings

The following section provides a discussion of the observations resulting from the analyses of both aims as they pertain to each of the FMT donor characteristics selected *a priori* based on a review of the existing literature. The discussion is organized by the statistical regression approach and variables utilized within those analyses, which were pre-identified in the theoretical framework, namely host, environmental, and microbiome features.

Analysis of Donor Characteristics Predictive of Passing

The purpose of the logistic regression utilized in the first aim was to explore and identify donor characteristics that were predictive of passing the FMT prospective donor screening program. Significant results from this analysis would be used to inform clinical and nursing practice on methods to help identify healthy donors earlier during the prospective donor

screening process. Eight variables were evaluated to determine whether the prospective consideration of these donor characteristics by nurses and other clinicians would reduce deferral rates and improve program efficiency. While the final model was significant at the 0.05 level, most of the variables considered for inclusion were found not to be predictive of the outcome. Specifically, gender, age, BMI, diet, alcohol and tobacco use, and seasonality were not predictive of whether a prospective donor passed or was deferred upon completion of the donor screening program. Only one donor related variable, frequency of bowel movements, was observed to be significant at the $p < 0.05$ level (p value of 0.018) and was included in the final model. The details of the results for each variable and the related implications for the stool donor screening process are discussed in the following section.

Host Features.

Gender. OpenBiome's FMT donor screening program encouraged participation by both genders. Despite this policy, 68.1% of prospective stool donors used in this study were male. This finding was notably different from trends in gender participation observed for blood donation, which report more equal representation (Greinacher, Fendrich, Brzenska, Kiefel, & Hoffmann, 2011). Analysis of the data demonstrated that the overall pass rate reported by male stool donors (18.9%) was marginally higher than for female stool donors (16.0%). Results from the logistic regression analysis confirmed that there was no relevant predictive relationship between gender and passing the prospective stool donor screening. This finding appeared to be contrary to existing literature that suggests women may be diagnosed with gastrointestinal disorders at higher rates and possibly more likely to be deferred (Bakken et al., 2011; Houghton et al., 2016; Lo et al., 2003). However, the results of this study did not support the assumption that gender predicts stool donor health or FMT suitability. Rather, the finding suggests that

prospective stool donor screening programs should remain agnostic as to gender and support equal participation by both genders in stool donation. Various factors may have contributed to the imbalance in gender participation observed in this study. Research to evaluate potential influences on individual gender motivations for donating stool could provide strategies for increasing participation by female stool donors.

Age. The age range for inclusion in OpenBiome's stool donor screening program was limited to donors between the ages of 18 and 50 years. Review of the study data demonstrated a progressive decrease in the pass rate as donor age increased, consistent with observations supported by the literature that suggest younger adults are generally healthier (Niccoli & Partridge, 2012). As such, the data suggested that younger adults may be more likely to pass the screening and may be less likely to be diagnosed with disorders that are generally found to increase in prevalence as one ages, such as high blood pressure and metabolic syndrome (Niccoli & Partridge, 2012). Results from the logistic regression analysis however, did not support age as predictive of a donor passing the screening. Rather, the study finding suggested that overall health for adults between the ages of 18 and 50 years remained relatively stable. These findings concur with results reported in the literature from studies evaluating the intestinal microbiome in healthy younger to middle aged adults (Odamaki et al., 2016). As such, this study did not support the need to further restrict the range in age for inclusion that is currently utilized by stool banks. Rather, it suggested that research to evaluate whether the age range can be broadened beyond the existing recommendation of 18 to 50 is warranted. This research may lead to greater access and a larger pool of suitable stool donors, which currently are deferred based on age alone.

Body Mass Index. Seminal research published in the past decade suggests a link between obesity, as measured by one's BMI, and the intestinal microbiome composition (Ley et al., 2005;

Turnbaugh et al., 2008; Turnbaugh et al., 2009). More recently, anecdotal case reports of post-FMT weight gain in recipients receiving material from obese donors were similarly reported (Alang & Kelly, 2015). For this reason, prospective stool donors with a BMI in the obese range of 30.0 kg/m² or greater are deferred from donating. Excess weight has been found to be associated with an increased risk and prevalence for chronic conditions, many of which are considered to be reasons for donor deferral i.e., metabolic syndrome and certain cancers (Guh et al., 2009; Kent et al., 2017). As such, this study sought to determine whether being in the overweight range of 25.0 kg/m² to 29.5 kg/m² affects prospective stool donors' general health and likelihood of being deferred. No correlation between an increase in BMI and donor deferral was found however. The results appear to support more recent literature that has introduced uncertainty regarding the link between stool donor BMI and post-procedure recipient weight gain (Fischer et al., 2018; Steevens, Roto, & DeCross, 2017), suggesting that factors other than donor BMI account for the development of obesity after FMT. As such, the results of this study did not support the need to change current BMI deferral recommendations. However, based on the conflicting results presented by this study and recent literature, re-evaluation of BMI as a metric for assessing stool donor health and suitability may be warranted. Further research would be useful to determine whether there are limitations in using BMI to evaluate stool donor health and suitability for donating fecal material. Such research could warrant the use of more precise clinical tools for identifying healthy FMT donor weight compositions and could result in improvements to the donor screening process.

Frequency of Bowel Movements. Most prospective donors included in this study sample reported daily bowel movements at a frequency considered by the literature and medical community to be within a healthy range of one to four (Connell, Hilton, Irvine, Lennard-Jones,

& Misiewicz, 1966; Heaton et al., 1992). However, findings from the logistic regression supported a predictive relationship between frequency of bowel movements and passing the stool donor screening. More specifically, the analysis revealed that donors with a frequency of bowel movements of two or more per day were more likely to be deferred during the screening, suggesting that there may be a possible connection between stool donor health and one's frequency of bowel movements. Although evidence for a correlation between intestinal microbiome composition and disease in asymptomatic individuals varies in the literature (Budree, Rao, et al., 2017; Vandeputte et al., 2016), further research on the impact of frequency of bowel movements and stool donor health may provide additional metrics to help nurses and other clinicians more readily identify suitable stool donors for use in FMT.

Based on the findings of this study, stool donor frequency of bowel movements may be a suitable proxy for assessing stool donor suitability in healthy people. The current standard for assessing donor health based on stool quality presented in the literature is the Bristol stool scale, which depends on visual inspection of a donor's stool (Budree, Rao, et al., 2017). While the Bristol stool scale has been shown in the literature to be associated with unique intestinal microbiome compositions (Vandeputte et al., 2016), the ability of the scale to accurately assess stool efficacy for use in FMT is inconclusive (Budree, Rao, et al., 2017; Budree, Wong, et al., 2017). Further, visual inspection of donor stool to evaluate donor health occurs continuously throughout the donation period and only after a prospective donor has been cleared to donate for FMT use. Measuring donor frequency of bowel movements may be an easier approach for nurses and other clinicians to proactively evaluate donor health in comparison to the Bristol stool scale. Inquiring about one's frequency of bowel movements is easily measured via a survey and can be conducted by nurses and other clinicians earlier in the screening process. This may allow nurses

and other clinicians to prioritize donor screening and defer donors who are not suitable earlier in the process, thus reducing costs and time to screen prospective stool donors.

Environmental Features.

Diet. Most prospective donors included in this study (86.7%) reported consuming an animal-based diet. Despite reports in the literature that certain diets promote unique bacterial intestinal compositions (De Filippo et al., 2010; Wu et al., 2011; Zimmer et al., 2012), the literature is inconclusive when it comes to identifying an optimal diet for stool donors. Some studies suggest that individuals who consume meat may be at higher risk for certain diseases that meet deferral criteria in the donor screening process, perhaps due to decreased dietary fiber or consumption of undercooked meat (Aiken et al., 2013; J. Slavin, 2013; J. L. Slavin, 2008). As such, there is an assumption that these donors may be more likely to be deferred during the prospective donor screening process. However, this was not supported by the logistic regression results, which demonstrated that diet was not predictive of passing or being deferred from the stool donor screening program. As such, selecting donors based on diet was not supported by this study and may not be informative in determining stool donor health or suitability for donating stool. Rather, this finding suggested that an evaluation of the use of deferral criteria based on donor dietary preferences may be warranted. Evidence supporting the value of including healthy stool donors despite their consumption of less common but increasingly popular fad diets, such as paleo or low carbohydrate diets, may increase the availability of healthy stool donors and the overall pass rate for prospective stool donor screening programs.

Alcohol Use. Current literature is lacking in evidence to support a relationship between low to moderate alcohol intake on donor health and suitability for stool donation. As such, there is value in evaluating whether regular low to moderate alcohol consumption is predictive of stool

donor health and suitability as a first step toward generating knowledge in this area. Although most prospective donors reported no alcohol consumption, the results of this study indicated that there was no relationship between the level of alcohol intake and passage of the donor screening. This finding suggests that low to moderate alcohol consumption may not have the same impact on donor health and passing the prospective donor screening or on the intestinal microbiome composition as has been observed in research on long-term, heavy alcohol use (Bode et al., 1984; P. Chen & Schnabl, 2014; Y. Chen et al., 2011; Mutlu et al., 2012). As such, the results of this study did not support the inclusion of alcohol consumption as a criterion for determining donor health and suitability for donating stool. However, given potential biases that can be introduced by donor self-report, the evaluation of other measures for quantifying the effects of regular, heavy alcohol consumption, such as binge drinking, on donor suitability may be warranted to ensure these donors continue to be identified and deferred.

Tobacco Use. Prospective donors who reported regular use of tobacco over the allowable threshold were deferred from donating stool due to the long-term effects of tobacco use on donor health as demonstrated by the literature (Giovannucci, 2001; Giovannucci & Martinez, 1996; N. Morgan, 1996). Infrequent and former tobacco users, however, were not deferred, citing a lack of research to support the deferral of this population. While the literature is limited in regard to the impact of infrequent tobacco use on the health and suitability of stool donors, there is some evidence to suggest that periodic and past tobacco use continue to negatively impact health for certain gastrointestinal related diseases long-term (Kuenzig et al., 2016). As such, evaluating donors for a possible dosage effect for tobacco use with passing the prospective donor screening program was warranted. Results of the logistic regression, however, demonstrated that there was no predictive difference in the pass rates for donors regardless of whether they reported zero,

past, or infrequent tobacco use, suggesting that stool donor healthy and suitability were not influenced by the reported levels of donor tobacco use. This finding provided further evidence to support the research of Tu and colleagues (2017), who reported that there was no difference in the intestinal microbiome of non-smoking donors and donors who smoked occasionally. As such, this evidence supported the continued use of current screening guidelines to evaluate tobacco use and donor suitability.

Seasonality. Evaluation of prospective donors occurred throughout the year with the majority of donors screened during the fall and winter months (29.2% per season). Trends in seasonality are not factored into OpenBiome's prospective donor screening program. However, the rates of donor deferral resulting from infectious diseases such as influenza and rotavirus may coincide with the peak incidence reported in the literature during the fall and winter months (Fisman, 2007). As such, this study aimed to assess whether there would be a similar uptick in donor deferral during the fall and winter months. Despite the sources to support this assumption in the literature, deferral rates across the seasons were similar and seasonality was found not to be predictive of donor deferral in the logistic regression analysis. This finding appeared to be in contrast to what has been reported in the literature and trends observed with the incidence for infectious disease. However, because no particular season was identified as predictive or optimal for screening donors, this study did not support consideration of seasonality during the evaluation of prospective donors nor that changes be made as to the frequency of donors screened throughout the year.

Impact of Alpha Diversity on Clinical Efficacy

The purpose of the second aim of this study was to assess whether donor 16S rRNA microbiome composition, as measured by Shannon's alpha diversity, was correlated with FMT

clinical efficacy in the treatment of patients with recurrent CDI. Results of the simple regression were used to evaluate the utility of including alpha diversity as a screening tool during the FMT donor screening process. Most studies to date have focused on evaluating the recipient's fecal microbiome to address engraftment of the donor material (Kelly et al., 2016; Seekatz et al., 2014; van Nood, Dijkgraaf, et al., 2013). Taking a different approach, this study evaluated the 16S rRNA intestinal microbiome of stool donor material for efficacy in treating patients with recurrent CDI. Available data about stool donor intestinal microbiome compositions reported as part of the aforementioned research literature were used to evaluate the findings of this study.

Microbiome Features. While the donor intestinal microbiome composition from this sample demonstrated some variability among individual donors, overall the sample cohort was comprised of generally healthy donors with diverse microbiome compositions consistent with a healthy gut (Cresci & Bawden, 2015; Dave, Higgins, Middha, & Rioux, 2012). Bacteria from nine different phyla predominantly colonize the healthy human gut (Gillilland, Young, & Huffnagle, 2012). Phyla are top level operational taxonomic units (OTU) representing the major lineages of bacteria and are comprised of numerous unique bacterial species. For this study's sample, bacteria from the *Firmicutes* and *Bacteroidetes* phyla comprised 91.4% of the total sequenced 16S rRNA microbiome data utilized in the analysis. In healthy adults, bacteria from these two aforementioned phyla have been shown to represent around 90% or greater of the bacterial composition of the intestinal microbiome (Rajilic-Stojanovic, Smidt, & de Vos, 2007). As a group, the intestinal microbiome composition sequenced for the entire sample cohort was consistent with observations found in the literature on the healthy human intestinal microbiome (Human Microbiome Project Consortium, 2012). This trend was similarly reflected at the individual donor level, with 55 (63.2%) of the donors having a 90% or greater intestinal

microbiome composition comprised of bacteria from the *Firmicutes* and *Bacteroidetes* phyla and most of the remaining donors (n=31, 35.6%) having at least 80%. Similar observations pertaining to bacterial richness were observed as part of the beta diversity analysis, further supporting the assumption that prospective donors who passed OpenBiome's stool donor screening program had an intestinal microbiome composition consistent with that observed in healthy individuals as reported in the literature.

Alpha Diversity. Alpha diversity of the donor samples provided a numerical metric of the level of diversity observed at the bacterial species level for each donor's intestinal microbiome composition. While reference thresholds for the range of alpha diversity associated with a healthy stool donor intestinal microbiome composition are not currently available in the literature, the range observed by this study was consistent those reported in similar research conducted on patient recipients of FMT (Kelly et al., 2016; Seekatz et al., 2014). Analysis of the scatterplot of alpha diversity against FMT clinical efficacy using the sample cohort demonstrated a positive relationship between the two variables. This observation suggested that, while donors in the sample were considered to be otherwise healthy, a correlation might exist between the bacterial diversity of the donor's fecal material and increased likelihood of preventing CDI recurrence in recipients. Despite these observations, the simple linear regression was not significant at the $p < 0.05$ level. As such, it could not be concluded that alpha diversity was correlated with any observed change in clinical efficacy resulting from the use of the stool donor material. This finding was consistent with results from similar studies on smaller donor cohorts presented in the literature (Budree, Wong, et al., 2017). As such, the results of this study provided additional evidence to support conclusions made by prior research that the intestinal microbiome composition among donors who pass the stool donor screening is predominantly homogenous in

regard to alpha diversity and generally efficacious in preventing recurrent CDI. Although the results of this study did not support the evaluation of donor alpha diversity during prospective donor screening, further research on donor and non-donor related characteristics might provide greater insight into factors that more directly contribute to increased FMT clinical efficacy rates.

Implications for Practice

The current process for identifying and selecting healthy stool donors that are suitable for donating fecal material for use in FMT presents many challenges for nurses and other clinicians who are tasked with leading this process on behalf of their patients. Although the results of this study did not definitively support the routine deferral of donors who pass two or more bowel movements per day, the findings can be used by nurses and other clinicians to prioritize the screening of specific donor populations that are more likely to pass. Stool donor screening does not currently utilize information regarding bowel habits as part of the triage or decision-making process. Inquiring about bowel habits at the initial online prescreening process could lead to increased program efficiency by reducing the number of in-person donor screening appointments. As such, targeted recruitment may prove to be more cost-effective when performing the in-person donor assessment by increasing the overall pass rates. While the results did provide suggestions on first steps that can be taken to improve the program, additional research would be needed to further evaluate whether donor frequency of bowel movements could serve as a proxy for some of the questions asked during the in-person screening process and possibly lead to additional improvements in the program.

The ability to target a specific donor profile may lead to a reduction in the delay to patient treatment in those who chose to use a directed donor approach. Nurses and other clinicians could counsel patient recipients to inquire about daily bowel habits with family and

friends. This conversation may be perceived as less obtrusive for both patient recipients and prospective donors. As such, this approach may be preferential to questioning prospective donors about other more private health related factors that increase one's risks for infectious or transmissible diseases, such as use of alcohol or drugs, previous diagnoses, and sexual behavior. This new information may decrease the time it takes for a patient to identify a suitable directed donor without having to disclose personal information about his or her own health.

The current process for evaluating stool donors is complex, costly, and time consuming. Thus, there is increased interest from the healthcare community for a rapid stool-based test to assess donor microbiome composition at the onset of screening to reduce delays, risks for transmitting an infectious organism, and increase clinical efficacy. Although the results of the analysis did not support the utility of incorporating alpha diversity as a biomarker during the stool donor screening process, the analysis indicated that there is value in further exploring the role of the intestinal microbiome in preventing CDI recurrence to identify different metrics and intestinal compositions that may be more efficacious.

Limitations

A shortfall of this study was the use of retrospective donor data for the analysis. The retrospective approach restricted the ability of the researcher to influence the way data about the variables were collected in a manner that would be most informative to the study. Several of the donor characteristics included in this study were presented as open-ended questions, for example diet and tobacco use. As such, this created an aspect of ambiguity in the answers and decision-making that may not have been consistently applied across the donors. Further, reliance on donor self-reporting introduced a potential bias. For example, a diet history collected with an accepted diet recall instrument may have been more informative and accurate than open-ended responses.

Ultimately, this study relied on data that may not have fully or accurately captured the donor characteristics, thus increasing the risk for insignificant findings.

The characteristics of the population from which the study sample was drawn also limit generalizability of the findings from this study. Evaluation of the non-significant prospective stool donor characteristics suggested that the overall population characteristics and study logistics might have influenced the sample demographics. In particular, the donor sample was heavily skewed toward younger individuals. Further, a notable percentage of donors were deferred due to logistics, despite being deemed healthy and suitable during the clinical assessment. These donors were unable to complete the entire screening process, which may have skewed the final pass or deferral outcome. The high density of young college students in the population from where the sample was drawn, and limitations imposed due to logistics of needing to travel often to the donation site suggest that caution be taken when applying these results to differing settings.

Recommendations for Future Research

The use of FMT to treat infectious and chronic diseases is a rapidly expanding field of medicine. However, the procedure is highly dependent on the availability of healthy stool donors and technological advancements. This study provided an evaluation of the health and microbiome characteristics of prospective stool donors in order to inform our understanding of how donor fecal material can be harnessed most efficiently and effectively by nurses and other clinicians to treat patients with recurrent CDI. As such, the results of this study serve as a platform for further inquiry in a program of research aimed at improving the outcomes of patients through the use of FMT from healthy stool donors. The success of this approach,

however, will depend on multidisciplinary research from the fields of nursing, microbiology, and medicine.

While this study identified donor frequency of bowel movements as predictive of passing the FMT donor screening program, additional opportunities exist to better understand the characteristics of optimal stool donors. This study did not evaluate deferral rationales to assess for any correlations with donor characteristics and behavior, which may be informative to the program. Other donor characteristics that were not captured by the prospective FMT donor screening program may have also contributed to the deferral of prospective donors. These include factors such as stress, exercise, and specific dietary habits, such as regular consumption of probiotics, dietary fiber, and vitamin D supplements. These may be informative to the screening process and donor outcomes. More research is needed to explore stool donor characteristics that were not captured by this study to evaluate whether these factors may or may not be predictive of passing the donor screening and could be utilized to improve the process.

Additional research is also needed to assess the value that each step in the lengthy and time-consuming prospective stool donor screening program provides to identifying optimal donors. Understanding donor characteristics that can serve as a proxy for infectious disease risk and a healthy microbiome composition, such as the frequency of bowel movements, could be utilized to shorten the steps during the clinical and laboratory screening. Evaluation of the frequency of bowel movements during screening could determine whether an infectious disease or microbiome-mediated risk was more likely to result in donor deferral. This information could be used to distinguish between real versus perceived risks and further build the literature and rationale for utilizing or removing barriers to stool donor screening in order to streamline the process while also maintain patient safety.

This study was not able to evaluate the stool donor characteristics in relation to intestinal microbiome composition. Though in-depth analysis of alpha diversity and donor characteristics was beyond the scope of this study, scatterplots of alpha diversity generated against donor frequency of bowel movements suggested trends in alpha diversity worth evaluating. This level of inquiry between additional intestinal microbiome markers with stool donor characteristics may be valuable in identifying whether specific donor characteristics are associated with unique and more efficacious intestinal microbiome compositions. This information may provide a platform for further study on the intestinal microbiome composition of stool donors and the clinical efficacy of FMT.

As an emerging area of science, there are many opportunities for future research on bacteria that colonize the intestines both at the phylum and species levels to better understand how they behave within different intestinal ecosystems. This research could generate information that assists nurses and other clinicians to expeditiously evaluate prospective donors to procure fecal material with consistently high rates of clinical efficacy and reduce further barriers to patient access. Despite the observation that bacteria from the *Firmicutes* and *Bacteroidetes* phyla predominantly comprise the intestinal microbiome of healthy donors, the efficacy of FMT material is likely to reside within the nuances of the species in the microbiome and how they interact with each other. Factors not evaluated as part of this study were the patient recipient comorbidities and medications and their influences on bacterial expression and engraftment. Understanding predictive factors for clinical efficacy that are external to the stool donors may provide a secondary approach to improving outcomes and increasing the likelihood of microbiome engraftment in the recipients.

Conclusions

The findings from this study were helpful in generating new knowledge and supporting existing literature on stool donor characteristics. These results were used to suggest changes that nurses and other clinicians could implement into the current process for screening stool donors to increase the pass rate, reduce costs and treatment delays, and potentially improve clinical efficacy. Results from the logistic regression identified donor reported frequency of bowel movements as informative in predicting whether a donor would be more likely to pass or be deferred. The findings of this study suggested that the ideal FMT donor candidates are donors who report passing fewer than two bowel movements per day on average. These donors may inherently have a lower risk for infectious or transmissible disease and possess a healthier microbiome. The remaining donor characteristics (gender, age, BMI, diet, tobacco and alcohol use, and seasonality) were not found to be predictive and thus supported maintaining the current criteria for screening stool donors. Similarly, the level of diversity observed by the stool donor intestinal microbiome was not found to be associated with a higher rate of FMT clinical efficacy. Despite the clinical interest in utilizing this tool during the prospective donor screening process, alpha diversity was not informative and thus not recommended as a potential biomarker for assessing the efficacy of donor stool or identifying more optimal stool donors. Additional research on factors not evaluated as part of this study could provide more evidence to better understand the factors that impact stool donor suitability and FMT clinical efficacy and lead to further improvements in the stool donor screening process.

References

Aiken, J., Bennett, R. W., Bhaduri, S., Bhatnagar, D., Bhunia, A. K., Bianchini, A., . . . Zygmunt, M. S. (2013). Guide to Foodborne Pathogens (Second ed.). Hoboken, NJ Wiley Blackwell.

Alang, N., & Kelly, C. R. (2015). Weight gain after fecal microbiota transplantation. Open Forum Infect Dis, 2(1), ofv004. doi:10.1093/ofid/ofv004

Bafeta, A., Yavchitz, A., Riveros, C., Batista, R., & Ravaud, P. (2017). Methods and reporting studies assessing fecal microbiota transplantation: A systematic review. Annals of Internal Medicine, 167(1), 34-39. doi:10.7326/m16-2810

Bakken, J. S., Borody, T., Brandt, L. J., Brill, J. V., Demarco, D. C., Franzos, M. A., . . . Surawicz, C. (2011). Treating *Clostridium difficile* infection with fecal microbiota transplantation. Clinical Gastroenterology and Hepatology, 9(12), 1044-1049. doi:10.1016/j.cgh.2011.08.014

Bakken, J. S., Polgreen, P. M., Beekmann, S. E., Riedo, F. X., & Streit, J. A. (2013). Treatment approaches including fecal microbiota transplantation for recurrent *Clostridium difficile* infection (RCDI) among infectious disease physicians. Anaerobe, 24, 20-24. doi:10.1016/j.anaerobe.2013.08.007

Bani, M., & Giussani, B. (2010). Gender differences in giving blood: A review of the literature. Blood Transfusion. Trasfusione del Sangue, 8(4), 278-287. doi:10.2450/2010.0156-09

Bernet-Camard, M. F., Lievin, V., Brassart, D., Neeser, J. R., Servin, A. L., & Hudault, S. (1997). The human *Lactobacillus* acidophilus strain LA1 secretes a nonbacteriocin antibacterial substance(s) active in vitro and in vivo. Applied and Environmental Microbiology, 63(7), 2747-2753.

Biagi, E., Candela, M., Turroni, S., Garagnani, P., Franceschi, C., & Brigidi, P. (2013). Ageing and gut microbes: Perspectives for health maintenance and longevity. Pharmacological Research, 69(1), 11-20. doi:10.1016/j.phrs.2012.10.005

Biedermann, L., Zeitz, J., Mwinyi, J., Sutter-Minder, E., Rehman, A., Ott, S. J., . . . Rogler, G. (2013). Smoking cessation induces profound changes in the composition of the intestinal microbiota in humans. PloS One, 8(3), e59260. doi:10.1371/journal.pone.0059260

Bode, J. C., Bode, C., Heidelbach, R., Durr, H. K., & Martini, G. A. (1984). Jejunal microflora in patients with chronic alcohol abuse. Hepato-Gastroenterology, 31(1), 30-34.

Bonder, M. J., Tigchelaar, E. F., Cai, X., Trynka, G., Cenit, M. C., Hrdlickova, B., . . . Zhernakova, A. (2016). The influence of a short-term gluten-free diet on the human gut microbiome. Genome Medicine, 8(1), 45. doi:10.1186/s13073-016-0295-y

Borgia, G., Maraolo, A. E., Foggia, M., Buonomo, A. R., & Gentile, I. (2015). Fecal microbiota transplantation for *Clostridium difficile* infection: Back to the future. Expert Opinion on Biological Therapy, 15(7), 1001-1014. doi:10.1517/14712598.2015.1045872

Brandt, L. J., Aroniadis, O. C., Mellow, M., Kanatzar, A., Kelly, C., Park, T., . . . Surawicz, C. (2012). Long-term follow-up of colonoscopic fecal microbiota transplant for recurrent *Clostridium difficile* infection. American Journal of Gastroenterology, 107(7), 1079-1087. doi:10.1038/ajg.2012.60

Britton, R. A., & Young, V. B. (2014). Role of the intestinal microbiota in resistance to colonization by *Clostridium difficile*. Gastroenterology, 146(6), 1547-1553. doi:10.1053/j.gastro.2014.01.059

Broecker, F., Kube, M., Klumpp, J., Schuppler, M., Biedermann, L., Hecht, J., . . . Moelling, K. (2013). Analysis of the intestinal microbiome of a recovered *Clostridium difficile* patient after fecal transplantation. Digestion, 88(4), 243-251. doi:10.1159/000355955

Brook, I. (1999). Bacterial interference. Critical Reviews in Microbiology, 25(3), 155-172. doi:10.1080/10408419991299211

Brumbaugh, D. E., De Zoeten, E. F., Pyo-Twist, A., Fidanza, S., Hughes, S., Dolan, S. A., . . . Dominguez, S. R. (2017). An intragastric fecal microbiota transplantation program for treatment of recurrent *Clostridium difficile* in children is efficacious, safe, and inexpensive. Journal of Pediatrics. doi:10.1016/j.jpeds.2017.10.016

Budree, S., Rao, S., Allegretti, J. R., Fischer, M., Kelly, C. R., Smith, M., . . . Kassam, Z. (2017). The association of donor stool consistency by bristol stool scale on microbial profile and clinical outcomes of fecal microbiota transplantation in *Clostridium difficile* infection. Gastroenterology, 152(5), S630.

Budree, S., Wong, W. F., Tu, E., Rao, S., Allegretti, J. R., Fischer, M., . . . Kassam, Z. (2017). Do specific bacteria drive clinical cure in fecal microbiota transplantation for *Clostridium difficile* infection? Clinical, microbial and metabolomic characterization of universal FMT donors. Gastroenterology, 152(5), S349.

Burns, L. J., Dubois, N. E., Smith, M. B., Mendolia, G. M., Burgess, J., Edelstein, C., . . . Kassam, Z. (2015). 499 donor recruitment and eligibility for fecal microbiota transplantation: Results from an international public stool bank. Gastroenterology, 148(4), 1. doi:10.1016/S0016-5085(15)30331-0

Burns, N., & Grove, S. (2001). The practice of nursing research: Conduct, critique and utilization (4th ed.). Philadelphia, Pennsylvania, USA.: W.B. Saunders.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods, 13(7), 581-583. doi:10.1038/nmeth.3869

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., . . . Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences of the United States of America, 108 Suppl 1, 4516-4522. doi:10.1073/pnas.1000080107

Center for Biologics Evaluation and Research. (2007). Guidance for industry eligibility determination for donors of human cells, tissues, and cellular and tissue based products (HCT/Ps). Rockville, MD: U.S. Department of Health and Human Services.

Center for Biologics Evaluation and Research. (2013). Enforcement policy regarding investigational new drug requirements for use of fecal microbiota for transplantation to treat *Clostridium difficile* infection not responsive to standard therapies. Retrieved from https://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Vaccines/UCM361393.pdf.

Centers for Disease Control and Prevention. (2017a). Adult obesity causes and consequences. Retrieved from https://www.cdc.gov/obesity/adult/causes.html

Centers for Disease Control and Prevention. (2017b). National health interview survey general concepts. Retrieved from https://www.cdc.gov/nchs/nhis/tobacco/tobacco_glossary.htm

Centers for Disease Control and Prevention. (2017c). Smoking and tobacco use: Health effects. Retrieved from https://www.cdc.gov/tobacco/basic_information/health_effects/

Chan, R. C., Reid, G., Irvin, R. T., Bruce, A. W., & Costerton, J. W. (1985). Competitive exclusion of uropathogens from human uroepithelial cells by *Lactobacillus* whole cells and cell wall fragments. Infection and Immunity, 47(1), 84-89.

Chang, J. Y., Antonopoulos, D. A., Kalra, A., Tonelli, A., Khalife, W. T., Schmidt, T. M., & Young, V. B. (2008). Decreased diversity of the fecal Microbiome in recurrent *Clostridium difficile*-associated diarrhea. Journal of Infectious Diseases, 197(3), 435-438. doi:10.1086/525047

Chen, P., & Schnabl, B. (2014). Host-microbiome interactions in alcoholic liver disease. Gut Liver, 8(3), 237-241. doi:10.5009/gnl.2014.8.3.237

Chen, Y., Yang, F., Lu, H., Wang, B., Chen, Y., Lei, D., . . . Li, L. (2011). Characterization of fecal microbial communities in patients with liver cirrhosis. Hepatology, 54(2), 562-572. doi:10.1002/hep.24423

Christian, N., Whitaker, B. K., & Clay, K. (2015). Microbiomes: Unifying animal and plant systems through the lens of community ecology theory. Front Microbiology, 6, 869. doi:10.3389/fmicb.2015.00869

Cohen, S. H., Gerding, D. N., Johnson, S., Kelly, C. P., Loo, V. G., McDonald, L. C., . . . Wilcox, M. H. (2010). Clinical practice guidelines for *Clostridium difficile* infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). Infection Control and Hospital Epidemiology, 31(5), 431-455. doi:10.1086/651706

Connell, A. M., Hilton, C., Irvine, G., Lennard-Jones, J. E., & Misiewicz, J. J. (1966). Variation of bowel habit in two population samples. Proceedings of the Royal Society of Medicine, 59(1), 11-12.

Cook, S. M., Glass, R. I., LeBaron, C. W., & Ho, M. S. (1990). Global seasonality of rotavirus infections. Bulletin of the World Health Organization, 68(2), 171-177.

Corthesy, B., Gaskins, H. R., & Mercenier, A. (2007). Cross-talk between probiotic bacteria and the host immune system. Journal of Nutrition, 137(3 Suppl 2), 781s-790s.

Cosnes, J. (2004). Tobacco and IBD: Relevance in the understanding of disease mechanisms and clinical practice. Best Practice & Research: Clinical Gastroenterology, 18(3), 481-496. doi:10.1016/j.bpg.2003.12.003

Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J., & Relman, D. A. (2012). The application of ecological theory toward an understanding of the human microbiome. Science, 336(6086), 1255-1262. doi:10.1126/science.1224203

Cox, L. M., & Blaser, M. J. (2013). Pathways in microbe-induced obesity. Cell Metabolism, 17(6), 883-894. doi:10.1016/j.cmet.2013.05.004

Cresci, G. A., & Bawden, E. (2015). Gut microbiome: What we do and don't know. Nutrition in Clinical Practice, 30(6), 734-746. doi:10.1177/0884533615609899

Daniel, H., Gholami, A. M., Berry, D., Desmarchelier, C., Hahne, H., Loh, G., . . . Clavel, T. (2014). High-fat diet alters gut microbiota physiology in mice. ISME Journal, 8(2), 295-308. doi:10.1038/ismej.2013.155

Dave, M., Higgins, P. D., Middha, S., & Rioux, K. P. (2012). The human gut microbiome: current knowledge, challenges, and future directions. Translational Research: The Journal of Laboratory and Clinical Medicine, 160(4), 246-257. doi:10.1016/j.trsl.2012.05.003

Davenport, E. R., Mizrahi-Man, O., Michelini, K., Barreiro, L. B., Ober, C., & Gilad, Y. (2014). Seasonal variation in human gut microbiome composition. PloS One, 9(3), e90731. doi:10.1371/journal.pone.0090731

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., . . . Turnbaugh, P. J. (2014). Diet rapidly and reproducibly alters the human gut microbiome. Nature, 505(7484), 559-563. doi:10.1038/nature12820

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., . . . Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proceedings of the National Academy of Sciences of the United States of America, 107(33), 14691-14696. doi:10.1073/pnas.1005963107

Department of Health and Human Services. (2017). Title 21 Code of Federal Regulations - Food and Drugs. Washington, DC: Food & Drug Administration.

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., & Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proceedings of the National Academy of Sciences of the United States of America, 107(26), 11971-11975. doi:10.1073/pnas.1002601107

Dowell, S. F. (2001). Seasonal variation in host susceptibility and cycles of certain infectious diseases. Emerging Infectious Diseases, 7(3), 369-374. doi:10.3201/eid0703.010301

Dubberke, E. R., & Olsen, M. A. (2012). Burden of *Clostridium difficile* on the healthcare system. Clinical Infectious Diseases, 55 Suppl 2, S88-92. doi:10.1093/cid/cis335

Eiseman, B., Silen, W., Bascom, G. S., & Kauvar, A. J. (1958). Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis. Surgery, 44(5), 854-859.

Eloe-Fadrosh, E. A., & Rasko, D. A. (2013). The human microbiome: From symbiosis to pathogenesis. Annual Review of Medicine, 64, 145-163. doi:10.1146/annurev-med-010312-133513

Falagas, M. E., Rafailidis, P. I., & Makris, G. C. (2008). Bacterial interference for the prevention and treatment of infections. International Journal of Antimicrobial Agents, 31(6), 518-522. doi:10.1016/j.ijantimicag.2008.01.024

Field, A. (2009). Discovering statistics using SPSS (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Fischer, M., Kao, D., Kassam, Z., Smith, J., Louie, T., Sipe, B., . . . Allegretti, J. R. (2018). Stool donor body mass index does not affect recipient weight after a single fecal microbiota transplantation for *Clostridium difficile* infection. Clinical Gastroenterology and Hepatology, 16(8), 1351-1353.

Fischer, M., Kao, D., Mehta, S. R., Martin, T., Dimitry, J., Keshteli, A. H., . . . Kelly, C. R. (2016). Predictors of early failure after fecal microbiota transplantation for the therapy of *Clostridium difficile* infection: A multicenter study. American Journal of Gastroenterology, 111(7), 1024-1031. doi:10.1038/ajg.2016.180

Fisman, D. N. (2007). Seasonality of infectious diseases. Annual Review of Public Health, 28, 127-143. doi:10.1146/annurev.publhealth.28.021406.144128

Garcia-Gomez, E., Gonzalez-Pedrajo, B., & Camacho-Arroyo, I. (2013). Role of sex steroid hormones in bacterial-host interactions. BioMed Research International, 2013, 928290. doi:10.1155/2013/928290

Gillilland, M. G., Young, V. B., & Huffnagle, G. B. (2012). Gastrointestinal Microbial Ecology with Perspectives on Health and Disease. In L. R. Johnson, J. D. Kaunitz, H. M. Said, F. K. Ghishan, J. L. Merchant, & J. D. Wood (Eds.), Physiology of the Gastrointestinal Tract (5th ed., pp. 1119-1134). London, UK: Elsevier Inc.

Giovannucci, E. (2001). An updated review of the epidemiological evidence that cigarette smoking increases risk of colorectal cancer. Cancer Epidemiology, Biomarkers and Prevention, 10(7), 725-731.

Giovannucci, E., & Martinez, M. E. (1996). Tobacco, colorectal cancer, and adenomas: A review of the evidence. Journal of the National Cancer Institute, 88(23), 1717-1730.

Goodrich, J. K., Di Rienzi, S. C., Poole, A. C., Koren, O., Walters, W. A., Caporaso, J. G., . . . Ley, R. E. (2014). Conducting a microbiome study. Cell, 158(2), 250-262. doi:10.1016/j.cell.2014.06.037

Greinacher, A., Fendrich, K., Brzenska, R., Kiefel, V., & Hoffmann, W. (2011). Implications of demographics on future blood supply: a population-based cross-sectional study. Transfusion, 51(4), 702-709. doi:10.1111/j.1537-2995.2010.02882.x

Guh, D. P., Zhang, W., Bansback, N., Amarsi, Z., Birmingham, C. L., & Anis, A. H. (2009). The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. BMC Public Health, 9, 88. doi:10.1186/1471-2458-9-88

Guigoz, Y., Dore, J., & Schiffrin, E. J. (2008). The inflammatory status of old age can be nurtured from the intestinal environment. Current Opinion in Clinical Nutrition and Metabolic Care, 11(1), 13-20. doi:10.1097/MCO.0b013e3282f2bfdf

Gupta, A., & Khanna, S. (2014). Community-acquired *Clostridium difficile* infection: An increasing public health threat. Infection and Drug Resistance, 7, 63-72. doi:10.2147/idr.s46780

Haro, C., Rangel-Zuniga, O. A., Alcala-Diaz, J. F., Gomez-Delgado, F., Perez-Martinez, P., Delgado-Lista, J., . . . Camargo, A. (2016). Intestinal microbiota is influenced by gender and body mass index. PloS One, 11(5), e0154090. doi:10.1371/journal.pone.0154090

Heaton, K. W., Radvan, J., Cripps, H., Mountford, R. A., Braddon, F. E., & Hughes, A. O. (1992). Defecation frequency and timing, and stool form in the general population: a prospective study. Gut, 33(6), 818-824.

Helwick, C. (2015). Ideal donors for fecal transplant rare. Retrieved from http://www.medscape.com/viewarticle/844905

Hisada, T., Endoh, K., & Kuriki, K. (2015). Inter- and intra-individual variations in seasonal and daily stabilities of the human gut microbiota in Japanese. Archives of Microbiology, 197(7), 919-934. doi:10.1007/s00203-015-1125-0

Hopkins, M. J., Sharp, R., & Macfarlane, G. T. (2002). Variation in human intestinal microbiota with age. Digestive and Liver Disease, 34 Suppl 2, S12-18.

Houghton, L. A., Heitkemper, M., Crowell, M., Emmanuel, A., Halpert, A., McRoberts, J. A., & Toner, B. (2016). Age, gender and women's health and the patient. Gastroenterology. doi:10.1053/j.gastro.2016.02.017

Human Microbiome Project Consortium. (2012). Structure, function and diversity of the healthy human microbiome. Nature, 486(7402), 207-214. doi:10.1038/nature11234

Kapel, N., Thomas, M., Corcos, O., Mayeur, C., Barbot-Trystram, L., Bouhnik, Y., & Joly, F. (2014). Practical implementation of faecal transplantation. Clinical Microbiology and Infection, 20(11), 1098-1105. doi:10.1111/1469-0691.12796

Kassam, Z., Lee, C. H., Yuan, Y., & Hunt, R. H. (2013). Fecal microbiota transplantation for *Clostridium difficile* infection: Systematic review and meta-analysis. American Journal of Gastroenterology, 108(4), 500-508. doi:10.1038/ajg.2013.59

Kelly, C. R., de Leon, L., & Jasutkar, N. (2012). Fecal microbiota transplantation for relapsing *Clostridium difficile* infection in 26 patients: Methodology and results. Journal of Clinical Gastroenterology, 46(2), 145-149. doi:10.1097/MCG.0b013e318234570b

Kelly, C. R., Ihunnah, C., Fischer, M., Khoruts, A., Surawicz, C., Afzali, A., . . . Brandt, L. (2014). Fecal microbiota transplant for treatment of *Clostridium difficile* infection in immunocompromised patients. American Journal of Gastroenterology, 109(7), 1065-1071. doi:10.1038/ajg.2014.133

Kelly, C. R., Khoruts, A., Staley, C., Sadowsky, M. J., Abd, M., Alani, M., . . . Brandt, L. J. (2016). Effect of Fecal Microbiota Transplantation on Recurrence in Multiply Recurrent *Clostridium difficile* Infection: A Randomized Trial. Annals of Internal Medicine, 165(9), 609-616. doi:10.7326/m16-0271

Kent, S., Fusco, F., Gray, A., Jebb, S. A., Cairns, B. J., & Mihaylova, B. (2017). Body mass index and healthcare costs: a systematic literature review of individual participant data studies. Obesity Reviews, 18(8), 869-879. doi:10.1111/obr.12560

Khan, R., & Cheesbrough, J. (2003). Impact of changes in antibiotic policy on *Clostridium difficile*-associated diarrhoea (CDAD) over a five-year period in a district general hospital. Journal of Hospital Infection, 54(2), 104-108.

Khanna, S., Pardi, D. S., Aronson, S. L., Kammer, P. P., Orenstein, R., St Sauver, J. L., . . . Zinsmeister, A. R. (2012). The epidemiology of community-acquired *Clostridium difficile* infection: A population-based study. American Journal of Gastroenterology, 107(1), 89-95. doi:10.1038/ajg.2011.398

Khoruts, A., Dicksved, J., Jansson, J. K., & Sadowsky, M. J. (2010). Changes in the composition of the human fecal microbiome after bacteriotherapy for recurrent *Clostridium difficile*-associated diarrhea. Journal of Clinical Gastroenterology, 44(5), 354-360. doi:10.1097/MCG.0b013e3181c87e02

Khoruts, A., & Weingarden, A. R. (2014). Emergence of fecal microbiota transplantation as an approach to repair disrupted microbial gut ecology. Immunology Letters, 162(2 Pt A), 77-81. doi:10.1016/j.imlet.2014.07.016

Koren, O., Goodrich, J. K., Cullender, T. C., Spor, A., Laitinen, K., Backhed, H. K., . . . Ley, R. E. (2012). Host remodeling of the gut microbiome and metabolic changes during pregnancy. Cell, 150(3), 470-480. doi:10.1016/j.cell.2012.07.008

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and Environmental Microbiology, 79(17), 5112-5120. doi:10.1128/aem.01043-13

Kuczynski, J., Liu, Z., Lozupone, C., McDonald, D., Fierer, N., & Knight, R. (2010). Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. Nat Methods, 7(10), 813-819. doi:10.1038/nmeth.1499

Kuenzig, M. E., Lee, S. M., Eksteen, B., Seow, C. H., Barnabe, C., Panaccione, R., & Kaplan, G. G. (2016). Smoking influences the need for surgery in patients with the inflammatory bowel diseases: A systematic review and meta-analysis incorporating disease duration. BMC Gastroenterology, 16(1), 143. doi:10.1186/s12876-016-0555-8

Kump, P. K., Krause, R., Allerberger, F., & Hogenauer, C. (2014). Faecal microbiota transplantation--the Austrian approach. Clinical Microbiology and Infection, 20(11), 1106-1111. doi:10.1111/1469-0691.12801

Lessa, F. C., Winston, L. G., & McDonald, L. C. (2015). Burden of *Clostridium difficile* infection in the United States. New England Journal of Medicine, 372(24), 2369-2370. doi:10.1056/NEJMc1505190

Ley, R. E., Backhed, F., Turnbaugh, P., Lozupone, C. A., Knight, R. D., & Gordon, J. I. (2005). Obesity alters gut microbial ecology. Proceedings of the National Academy of Sciences of the United States of America, 102(31), 11070-11075. doi:10.1073/pnas.0504978102

Lo, W., Sano, K., Lebwohl, B., Diamond, B., & Green, P. H. (2003). Changing presentation of adult celiac disease. Digestive Diseases and Sciences, 48(2), 395-398.

Loo, V. G., Bourgault, A. M., Poirier, L., Lamothe, F., Michaud, S., Turgeon, N., . . . Dascal, A. (2011). Host and pathogen factors for *Clostridium difficile* infection and colonization. New England Journal of Medicine, 365(18), 1693-1703. doi:10.1056/NEJMoa1012413

Mariat, D., Firmesse, O., Levenez, F., Guimaraes, V., Sokol, H., Dore, J., . . . Furet, J. P. (2009). The *Firmicutes/Bacteroidetes* ratio of the human microbiota changes with age. BMC Microbiology, 9, 123. doi:10.1186/1471-2180-9-123

McFarland, L. V. (2012). The role of compassion in the practice of evidence-based medicine. American Journal of Gastroenterology, 107(5), 768-769. doi:10.1038/ajg.2011.489

McNabb-Baltar, J., Yaghoobi, M., O'Byrne, M., Soulellis, C. A., & Trinh, Q.-D. (2013). Burden of *Clostridium difficile*-associated hospitalizations in the United States. Gastroenterology, 144(5), S-239. doi:10.1016/S0016-5085(13)60846-X

Meighani, A., Hart, B. R., Mittal, C., Miller, N., John, A., & Ramesh, M. (2016). Predictors of fecal transplant failure. European Journal of Gastroenterology and Hepatology, 28(7), 826-830. doi:10.1097/meg.0000000000000614

Mokdad, A. H., Ford, E. S., Bowman, B. A., Dietz, W. H., Vinicor, F., Bales, V. S., & Marks, J. S. (2003). Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. Journal of the American Medical Association, 289(1), 76-79.

Monto, A. S. (2002). Epidemiology of viral respiratory infections. American Journal of Medicine, 112 Suppl 6A, 4s-12s.

Morgan, N. (1996). Strategies for colic. Birth Gazette, 12(4), 18-23.

Morgan, X. C., Tickle, T. L., Sokol, H., Gevers, D., Devaney, K. L., Ward, D. V., . . . Huttenhower, C. (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome Biology, 13(9), R79. doi:10.1186/gb-2012-13-9-r79

Mutlu, E. A., Gillevet, P. M., Rangwala, H., Sikaroodi, M., Naqvi, A., Engen, P. A., . . . Keshavarzian, A. (2012). Colonic microbiome is altered in alcoholism. American Journal of Physiology: Gastrointestinal and Liver Physiology, 302(9), G966-978. doi:10.1152/ajpgi.00380.2011

Myers, A. (2015). Fecal microbiota transplants: Overcoming the barriers of implementation. Infection Control Today. Retrieved from http://www.infectioncontroltoday.com/articles/2015/12/fecal-microbiota-transplants-overcoming-the-barriers-of-implementation.aspx

Nagao-Kitamoto, H., Kitamoto, S., Kuffa, P., & Kamada, N. (2016). Pathogenic role of the gut microbiota in gastrointestinal diseases. Intestinal Research, 14(2), 127-138. doi:10.5217/ir.2016.14.2.127

Niccoli, T., & Partridge, L. (2012). Ageing as a risk factor for disease. Current Biology, 22(17), R741-752. doi:10.1016/j.cub.2012.07.024

Nuttall, F. Q. (2015). Body mass index: Obesity, BMI, and health: A critical review. Nutrition Today, 50(3), 117-128. doi:10.1097/nt.0000000000000092

O'Sullivan, A., He, X., McNiven, E. M., Haggarty, N. W., Lonnerdal, B., & Slupsky, C. M. (2013). Early diet impacts infant rhesus gut microbiome, immunity, and metabolism. Journal of Proteome Research, 12(6), 2833-2845. doi:10.1021/pr4001702

O'Toole, P. W., & Jeffery, I. B. (2015). Gut microbiota and aging. Science, 350(6265), 1214-1215. doi:10.1126/science.aac8469

Odamaki, T., Kato, K., Sugahara, H., Hashikura, N., Takahashi, S., Xiao, J. Z., . . . Osawa, R. (2016). Age-related changes in gut microbiota composition from newborn to centenarian: A cross-sectional study. BMC Microbiology, 16, 90. doi:10.1186/s12866-016-0708-5

Osman, M., Stoltzner, Z., O'Brien, K., Ling, K., Koelsch, E., Dubois, N., . . . Kassam, Z. (2016). Donor efficacy in fecal microbiota transplantation for recurrent *Clostridium difficile*: Evidence from a 1,999-patient cohort. Open Forum Infectious Diseases, 3(suppl_1), 841-841. doi:10.1093/ofid/ofw194.48

Ozaki, E., Kato, H., Kita, H., Karasawa, T., Maegawa, T., Koino, Y., . . . Nakamura, S. (2004). *Clostridium difficile* colonization in healthy adults: transient colonization and correlation with enterococcal colonization. Journal of Medical Microbiology, 53(Pt 2), 167-172. doi:10.1099/jmm.0.05376-0

Pakyz, A. L., Moczygemba, L. R., VanderWielen, L. M., & Edmond, M. B. (2016). Fecal microbiota transplantation for recurrent *Clostridium difficile* infection: The patient experience. American Journal of Infection Control, 44(5), 554-559. doi:10.1016/j.ajic.2016.01.018

Paramsothy, S., Borody, T. J., Lin, E., Finlayson, S., Walsh, A. J., Samuel, D., . . . Kamm, M. A. (2015). Donor recruitment for fecal microbiota transplantation. Inflammatory Bowel Diseases, 21(7), 1600-1606. doi:10.1097/mib.0000000000000405

Penders, J., Stobberingh, E. E., van den Brandt, P. A., & Thijs, C. (2007). The role of the intestinal microbiota in the development of atopic disorders. Allergy, 62(11), 1223-1236. doi:10.1111/j.1398-9995.2007.01462.x

Peterson, D. A., Frank, D. N., Pace, N. R., & Gordon, J. I. (2008). Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. Cell Host Microbe, 3(6), 417-427. doi:10.1016/j.chom.2008.05.001

Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J. A., . . . Guyer, M. (2009). The National Institute of Health Human Microbiome Project. Genome Research, 19(12), 2317-2323. doi:10.1101/gr.096651.109

Petrof, E. O., & Khoruts, A. (2014). From stool transplants to next-generation microbiota therapeutics. Gastroenterology, 146(6), 1573-1582. doi:10.1053/j.gastro.2014.01.004

Pickering, L. K., Bartlett, A. V., 3rd, Reves, R. R., & Morrow, A. (1988). Asymptomatic excretion of rotavirus before and after rotavirus diarrhea in children in day care centers. Journal of Pediatrics, 112(3), 361-365.

Plitchta, S. B., & Kelvin, E. A. (2013). Munro's statistical methods for health care research (6th ed.). Philadelphia: Wolters Kluwer.

QIIME 2 development team. (2018). QIIME2 (Version 2018.8). Retrieved from https://qiime2.org/

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., . . . Wang, J. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature, 490(7418), 55-60. doi:10.1038/nature11450

Rajilic-Stojanovic, M., Smidt, H., & de Vos, W. M. (2007). Diversity of the human gastrointestinal tract microbiota revisited. Environ Microbiol, 9(9), 2125-2136. doi:10.1111/j.1462-2920.2007.01369.x

Ray, A., & Jones, C. (2016). Does the donor matter? Donor versus patient effects in the outcome of a next-generation microbiota-based drug trial for recurrent *Clostridium difficile* infection. Future Microbiology, 11, 611-616. doi:10.2217/fmb.16.10

Rea, M. C., O'Sullivan, O., Shanahan, F., O'Toole, P. W., Stanton, C., Ross, R. P., & Hill, C. (2012). *Clostridium difficile* carriage in elderly subjects and associated changes in the intestinal microbiota. Journal of Clinical Microbiology, 50(3), 867-875. doi:10.1128/jcm.05176-11

Samuel, B. P., Crumb, T. L., & Duba, M. M. (2014). What nurses need to know about fecal microbiota transplantation: Education, assessment, and care for children and young adults. Journal of Pediatric Nursing, 29(4), 354-361. doi:10.1016/j.pedn.2014.01.013

Schubert, A. M., Rogers, M. A., Ring, C., Mogle, J., Petrosino, J. P., Young, V. B., . . . Schloss, P. D. (2014). Microbiome data distinguish patients with *Clostridium difficile* infection and non-C. difficile-associated diarrhea from healthy controls. MBio, 5(3), e01021-01014. doi:10.1128/mBio.01021-14

Seekatz, A. M., Aas, J., Gessert, C. E., Rubin, T. A., Saman, D. M., Bakken, J. S., & Young, V. B. (2014). Recovery of the gut microbiome following fecal microbiota transplantation. MBio, 5(3), e00893-00814. doi:10.1128/mBio.00893-14

Sha, S., Liang, J., Chen, M., Xu, B., Liang, C., Wei, N., & Wu, K. (2014). Systematic review: Faecal microbiota transplantation therapy for digestive and nondigestive disorders in adults and children. Alimentary Pharmacology and Therapeutics, 39(10), 1003-1032. doi:10.1111/apt.12699

Shannon, C. E. (1997). The mathematical theory of communication. 1963. MD Computing, 14(4), 306-317.

Slavin, J. (2013). Fiber and prebiotics: mechanisms and health benefits. Nutrients, 5(4), 1417-1435. doi:10.3390/nu5041417

Slavin, J. L. (2008). Position of the American Dietetic Association: health implications of dietary fiber. Journal of the American Dietetic Association, 108(10), 1716-1731.

Smith, M. B., Kelly, C., & Alm, E. J. (2014). Policy: How to regulate faecal transplants. Nature, 506(7488), 290-291.

Steevens, C. D., Roto, D., & DeCross, A. J. (2017). Obese stool donors in fecal microbiota transplantation: Not associated with recipient weight gain! Gastroenterology, 152(5), S1007 - S1008.

Stranges, P. M., Hutton, D. W., & Collins, C. D. (2013). Cost-effectiveness analysis evaluating fidaxomicin versus oral vancomycin for the treatment of *Clostridium difficile* infection in the United States. Value in Health, 16(2), 297-304. doi:10.1016/j.jval.2012.11.004

Surawicz, C. M., Brandt, L. J., Binion, D. G., Ananthakrishnan, A. N., Curry, S. R., Gilligan, P. H., . . . Zuckerbraun, B. S. (2013). Guidelines for diagnosis, treatment, and prevention of *Clostridium difficile* infections. American Journal of Gastroenterology, 108(4), 478-498; quiz 499. doi:10.1038/ajg.2013.4

Tigchelaar, E. F., Bonder, M. J., Jankipersadsing, S. A., Fu, J., Wijmenga, C., & Zhernakova, A. (2016). Gut microbiota composition associated with stool consistency. Gut, 65(3), 540-542. doi:10.1136/gutjnl-2015-310328

Tourneur, E., & Chassin, C. (2013). Neonatal immune adaptation of the gut and its role during infections. Clinical & Developmental Immunology, 2013, 270301. doi:10.1155/2013/270301

Tu, E., Osman, M., Leith, T., Rao, S., Ling, K., Smith, M., . . . Budree, S. (2017). The impact of donor diet and lifestyle habits on microbial profile and clinical outcomes of fecal microbiota transplantation in *Clostridium difficile* infection. Paper presented at the Massachusetts Institute of Technology - Harvard Microbiome Symposium Cambridge, MA.

Tuddenham, S., & Sears, C. L. (2015). The intestinal microbiome and health. Current Opinion in Infectious Diseases, 28(5), 464-470. doi:10.1097/qco.0000000000000196

Turnbaugh, P. J., Backhed, F., Fulton, L., & Gordon, J. I. (2008). Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. Cell Host Microbe, 3(4), 213-223. doi:10.1016/j.chom.2008.02.015

Turnbaugh, P. J., & Gordon, J. I. (2009). The core gut microbiome, energy balance and obesity. Journal of Physiology, 587(Pt 17), 4153-4158. doi:10.1113/jphysiol.2009.174136

Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., . . . Gordon, J. I. (2009). A core gut microbiome in obese and lean twins. Nature, 457(7228), 480-484. doi:10.1038/nature07540

Tyler, A. D., Smith, M. I., & Silverberg, M. S. (2014). Analyzing the human microbiome: A "how to" guide for physicians. American Journal of Gastroenterology, 109(7), 983-993.

United States Census Bureau. (2011). Age and sex composition: 2010. Retrieved from https://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf

United States Department of Health and Human Services. (2015). 2015-2020 Edition of the Dietary Guidelines for Americans. Retrieved from https://health.gov/dietaryguidelines/2015/guidelines/table-of-contents/

van Nood, E., Dijkgraaf, M. G., & Keller, J. J. (2013). Duodenal infusion of feces for recurrent *Clostridium difficile*. New England Journal of Medicine, 368(22), 2145. doi:10.1056/NEJMc1303919

van Nood, E., Vrieze, A., Nieuwdorp, M., Fuentes, S., Zoetendal, E. G., de Vos, W. M., . . . Keller, J. J. (2013). Duodenal infusion of donor feces for recurrent *Clostridium difficile*. New England Journal of Medicine, 368(5), 407-415. doi:10.1056/NEJMoa1205037

Vandeputte, D., Falony, G., Vieira-Silva, S., Tito, R. Y., Joossens, M., & Raes, J. (2016). Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Gut, 65(1), 57-62. doi:10.1136/gutjnl-2015-309618

Vincent, C., & Manges, A. R. (2015). Antimicrobial use, human gut microbiota and *Clostridium difficile* colonization and infection. Antibiotics (Basel), 4(3), 230-253. doi:10.3390/antibiotics4030230

Weissman, J. S., & Coyle, W. (2012). Stool transplants: Ready for prime time? Current Gastroenterology Reports, 14(4), 313-316. doi:10.1007/s11894-012-0263-7

World Health Organization. (2012). Blood donor selection: Guidelines on assessing donor suitability for blood donation. Luxembourg: Author.

Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y. Y., Keilbaugh, S. A., . . . Lewis, J. D. (2011). Linking long-term dietary patterns with gut microbial enterotypes. Science, 334(6052), 105-108. doi:10.1126/science.1208344

Youngster, I., Sauk, J., Pindar, C., Wilson, R. G., Kaplan, J. L., Smith, M. B., . . . Hohmann, E. L. (2014). Fecal microbiota transplant for relapsing *Clostridium difficile* infection using a frozen inoculum from unrelated donors: a randomized, open-label, controlled pilot study. Clinical Infectious Diseases, 58(11), 1515-1522. doi:10.1093/cid/ciu135

Zhang, F., Luo, W., Shi, Y., Fan, Z., & Ji, G. (2012). Should we standardize the 1,700-year-old fecal microbiota transplantation? American Journal of Gastroenterology, 107(11), 1755; author reply p.1755-1756. doi:10.1038/ajg.2012.251

Zimmer, J., Lange, B., Frick, J. S., Sauer, H., Zimmermann, K., Schwiertz, A., . . . Enck, P. (2012). A vegan or vegetarian diet substantially alters the human colonic faecal microbiota. European Journal of Clinical Nutrition, 66(1), 53-60. doi:10.1038/ejcn.2011.141

Zou, S., Musavi, F., Notari, E. P., Rios, J. A., Trouern-Trend, J., & Fang, C. T. (2008). Donor deferral and resulting donor loss at the American Red Cross Blood Services, 2001 through 2006. Transfusion, 48(12), 2531-2539. doi:10.1111/j.1537-2995.2008.01903.x

Appendix A. Change in descriptive statistics for continuous variables pre- and post-listwise removal

		N	Range	Minimum	Maximum
Age	Pre-Removal	767	31.0	18.0	49.0
	Post-Removal	744	31.0	18.0	49.0
	Change	23	0.0	0.0	0.0
Body Mass Index	Pre-Removal	751	13.5	16.3	29.8
	Post-Removal	744	12.9	16.9	29.8
	Change	7	0.6	-0.6	0.0
Frequency of Bowel Movements	Pre-Removal	758	4.5	0.5	5.0
	Post-Removal	744	4.5	0.5	5.0
	Change	14	0.0	0.0	0.0
Alcohol Use	Pre-Removal	744	20.0	0.0	20.0
	Post-Removal	744	20.0	0.0	20.0
	Change	0	0.0	0.0	0.0

		Mean		Std. Deviation	Variance
		Statistic	Std. Error		
Age	Pre-Removal	28.1	0.22	6.13	37.63
	Post-Removal	28.1	0.23	6.19	38.25
	Change	0.0	0.00	-0.10	-0.63
Body Mass Index	Pre-Removal	23.6	0.10	2.71	7.32
	Post-Removal	23.6	0.10	2.69	7.22
	Change	-0.02	0.00	0.02	0.11
Frequency of Bowel Movements	Pre-Removal	1.7	0.02	0.67	0.45
	Post-Removal	1.7	0.02	0.67	0.45
	Change	0.0	0.00	0.00	-0.01
Alcohol Use	Pre-Removal	2.9	0.10	2.69	7.26
	Post-Removal	2.9	0.10	2.71	7.33
	Change	0.0	0.00	0.00	-0.07

Appendix B. Change in descriptive statistics for categorical variables pre- and post- listwise removal

			Frequency	%	Cumulative %
Gender	Male	Pre-Removal	521	67.7	67.7
		Post-Removal	507	68.1	68.1
		Change	14	-0.5	
	Female	Pre-Removal	249	32.3	100.0
		Post-Removal	237	31.9	100.0
		Change	12	0.5	
	Total	Pre-Removal	770		
		Post-Removal	744		
		Change	26	3.4	
Diet	Meat Based	Pre-Removal	668	86.8	86.8
		Post-Removal	645	86.7	86.7
		Change	23	0.1	
	No Meat	Pre-Removal	102	13.2	100.0
		Post-Removal	99	13.3	100.0
		Change	3	-0.1	
	Total	Pre-Removal	770		
		Post-Removal	744		
		Change	26	3.4	
Tobacco Use	Never	Pre-Removal	621	81.4	81.4
		Post-Removal	605	81.3	81.3
		Change	16	0.1	
	Former	Pre-Removal	89	11.7	93.1
		Post-Removal	87	11.7	93.0
		Change	2	0.0	
	Current	Pre-Removal	53	6.9	100.0
		Post-Removal	52	7.0	100.0
		Change	1	0.0	
	Subtotal	Pre-Removal	763		
		Post-Removal	744		
		Change	19	2.5	
	Missing	Pre-Removal	7	0.9	
		Post-Removal	0	0.0	
		Change	7	0.9	
	Total	Pre-Removal	770		
		Post-Removal	744		
		Change	26	3.4	

			Frequency	%	Cumulative %
Seasonality	Winter	Pre-Removal	226	29.4	29.4
		Post-Removal	217	29.2	29.2
		Change	9	0.2	
	Spring	Pre-Removal	140	18.2	47.5
		Post-Removal	139	18.7	47.8
		Change	1	-0.5	
	Summer	Pre-Removal	178	23.1	70.6
		Post-Removal	171	23.0	70.8
		Change	7	0.1	
	Fall	Pre-Removal	226	29.4	100.0
		Post-Removal	217	29.2	100.0
		Change	9	0.2	
	Total	Pre-Removal	770		
		Post-Removal	744		
		Change	26	3.4	

Appendix C. Clinical assessment deferral categories utilized to assess prospective donor qualifications

Atopy, Asthma, or Allergies. Donors deferred for a positive diagnosis of two or more conditions related to atopy, asthma, or allergies e.g., atopic dermatitis, drug allergies, asthma, etc.

Autoimmune Conditions. Donors deferred for a positive history of an autoimmune related diagnosis e.g., psoriasis, or other recurrent skin condition.

Cancer History. Donors deferred for a positive history of a cancer diagnosis e.g., colon cancer.

Cardiovascular and Metabolic Conditions. Donors deferred for a positive diagnostic history of a cardiovascular of metabolic condition e.g., hypertension, diabetes type 1 or 2.

Current Health Status. Donors deferred if symptomatic for gastrointestinal or respiratory related illness during the screening or in the two weeks prior.

Diet and Exercise. Donors deferred for reporting a fad diet e.g., gluten-free, recent elimination diet, or regular avoidance of a specific food due to food intolerance.

Family History. Donors deferred for reporting a positive family history of a gastrointestinal related diagnosis where a risk for inheritance is suspected e.g., inflammatory bowel disease.

Gastrointestinal Conditions. Donors deferred for a positive history of gastrointestinal diagnosis e.g., inflammatory bowel disease, or for consulting with a gastroenterologist for treatment of a recurrent gastrointestinal symptom or gastrointestinal related medical concern.

Infectious Disease History. Donors deferred for a positive history of a diagnosis with a chronic communicable disease e.g., Human immunodeficiency virus or recent exposure that placed the donor at increased risk for contracting an infectious or communicable disease e.g., recent tick bite.

Medications and Supplements. Donors deferred for reporting regular use of a prescription or over-the-counter medication e.g., laxatives, antibiotics, etc., regular use of a supplement suspected of impacting the intestinal microbiome e.g., probiotics, or recent immunization with a vaccination utilizing an attenuated organism.

Mental Health and Well-Being. Donors deferred for a positive history of a mental health related diagnosis related e.g., generalized anxiety.

Musculoskeletal Conditions. Donors deferred for reporting a positive history of a musculoskeletal diagnosis that results in chronic pain e.g., fibromyalgia.

Neurological Conditions. Donors deferred for a positive history of a neurological diagnosis e.g., multiple sclerosis.

Sexual Behavior and History. Donors deferred for reporting recent sexual behavior associated with an increased risk for infectious or communicable disease transmitted through sexual contact e.g., anonymous sexual contact.

Social History. Donors deferred for reporting active or regular volunteer work in an environment at high risk for exposure to an infectious or communicable disease e.g., hospital workers, use of illicit drugs, or recent tattoo or piercing.

Surgical and Other Medical History. Donors deferred for major surgeries e.g., transplantation or other procedures or medical diagnosis requiring frequent or increased hospital exposure.

Travel History. Donors deferred for recent travel to areas considered high risk for infectious or communicable disease.

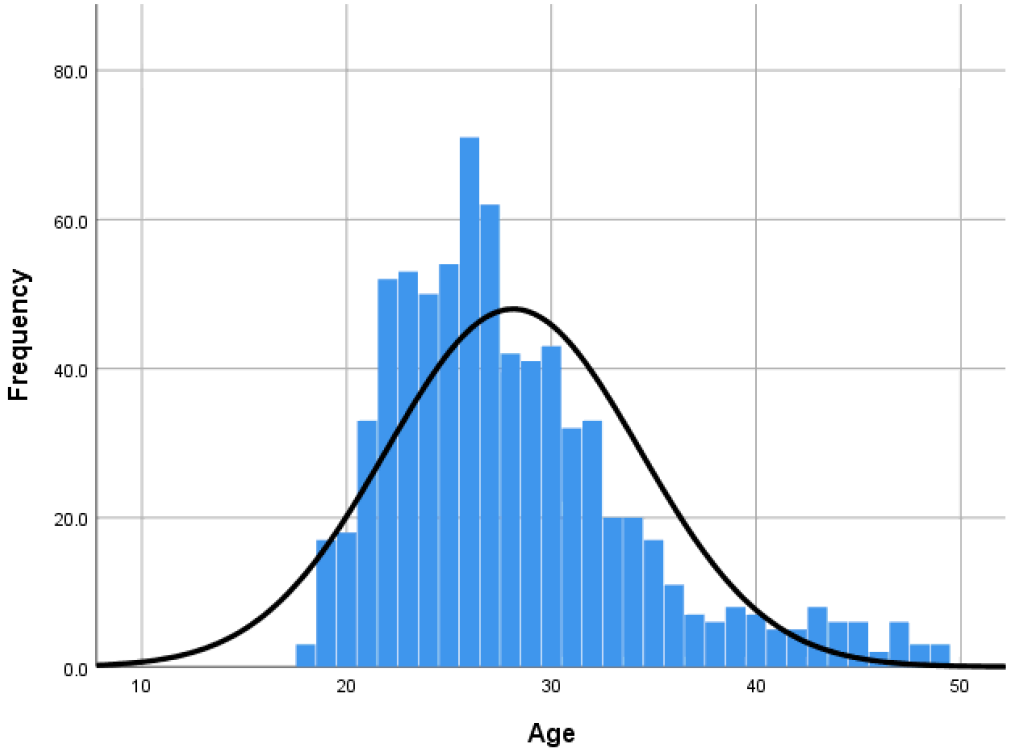
Vital Signs. Donors deferred for a finding of abnormal vital signs e.g., repeated hypertensive measurements, Body Mass Index greater than or equal to 30, etc.

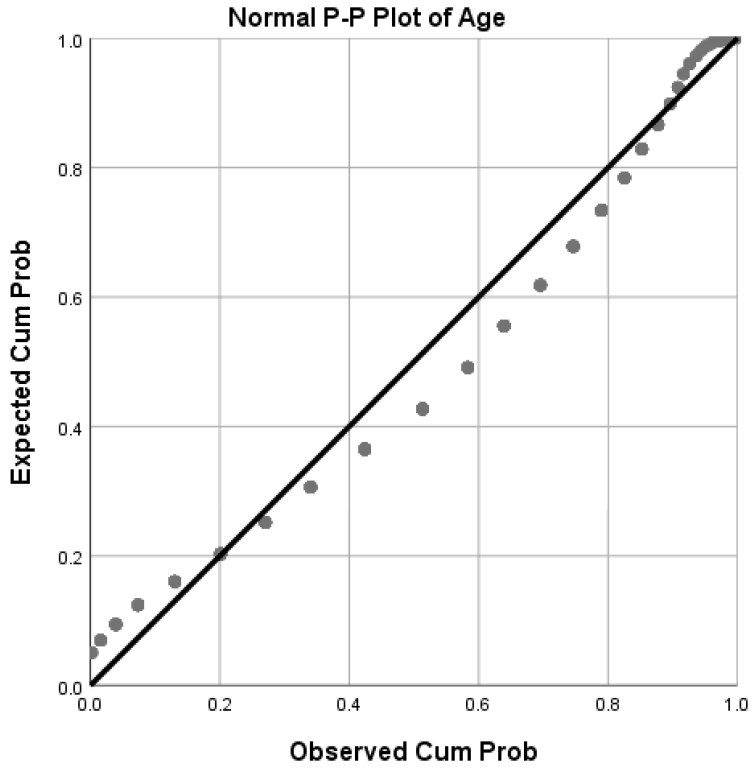
Women's Health. Donors deferred for being pregnant or actively breast-feeding.

Other. Donors deferred for logistical reasons or other, not previously reported infectious or communicable diagnosis or medical conditions that may impact the intestinal microbiome.

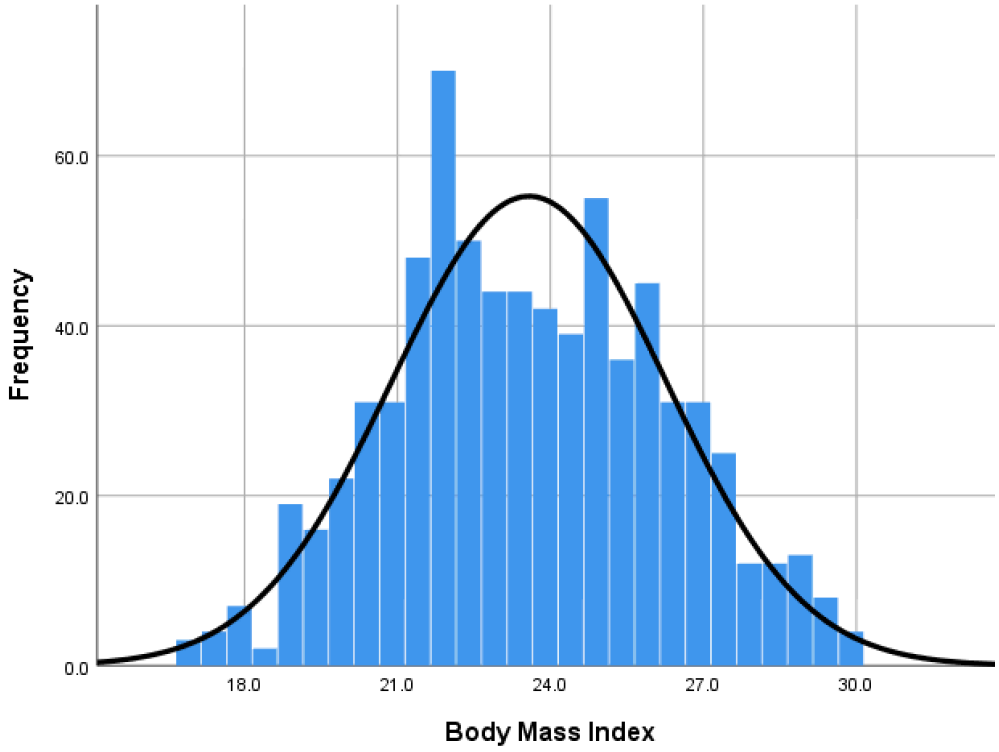
Appendix D. Histograms and P-P plots for continuous variables (age, body mass index, frequency of bowel movements, and alcohol use)

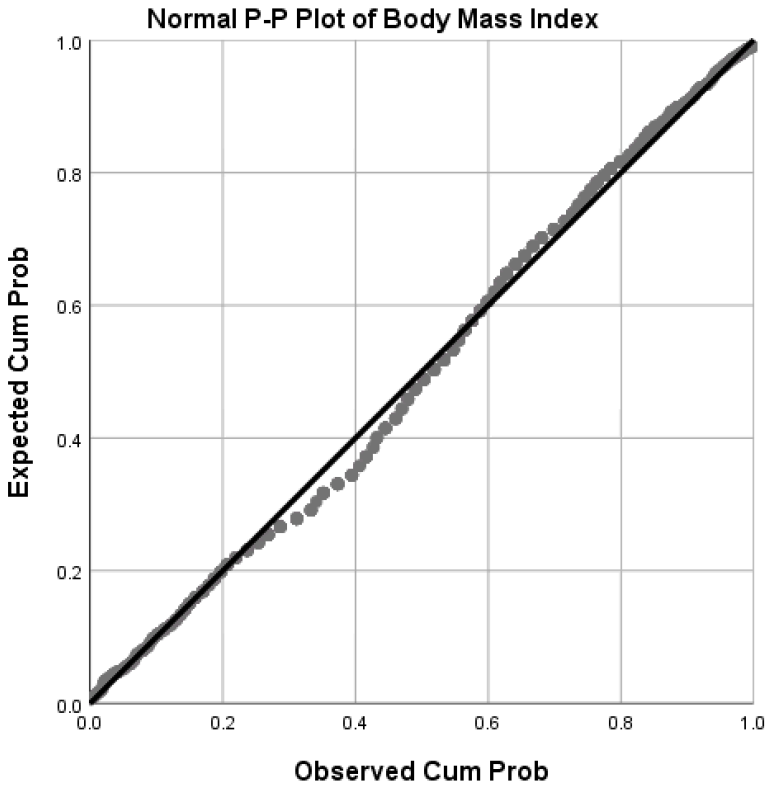
Histogram of Age



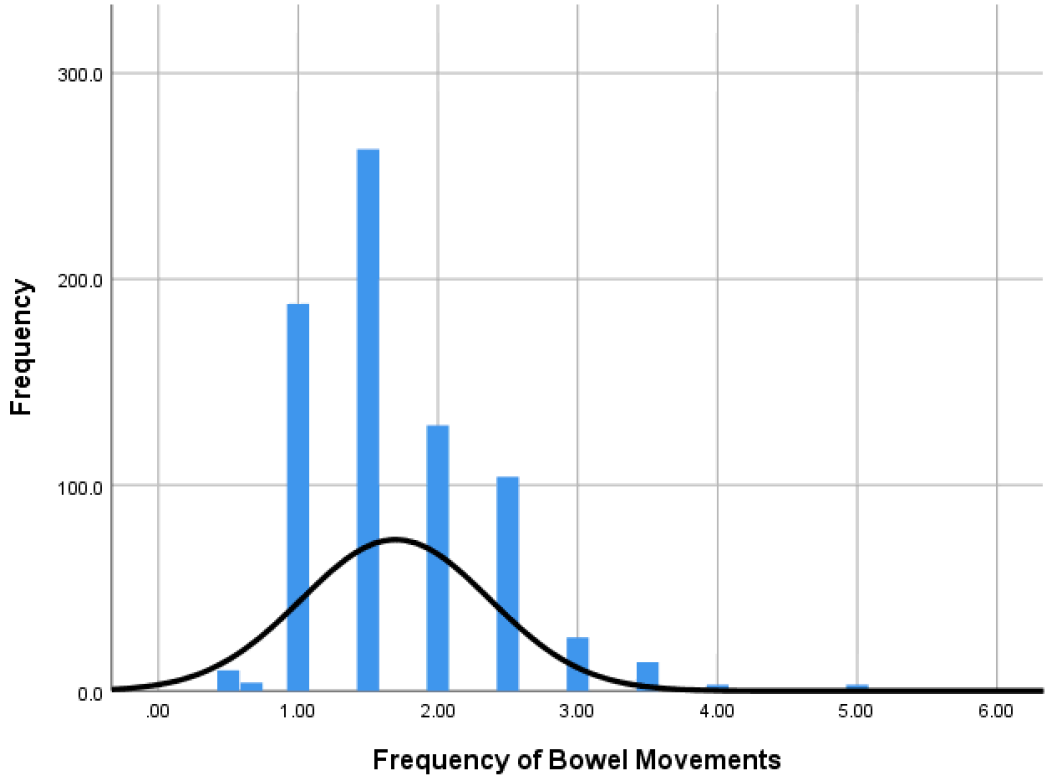


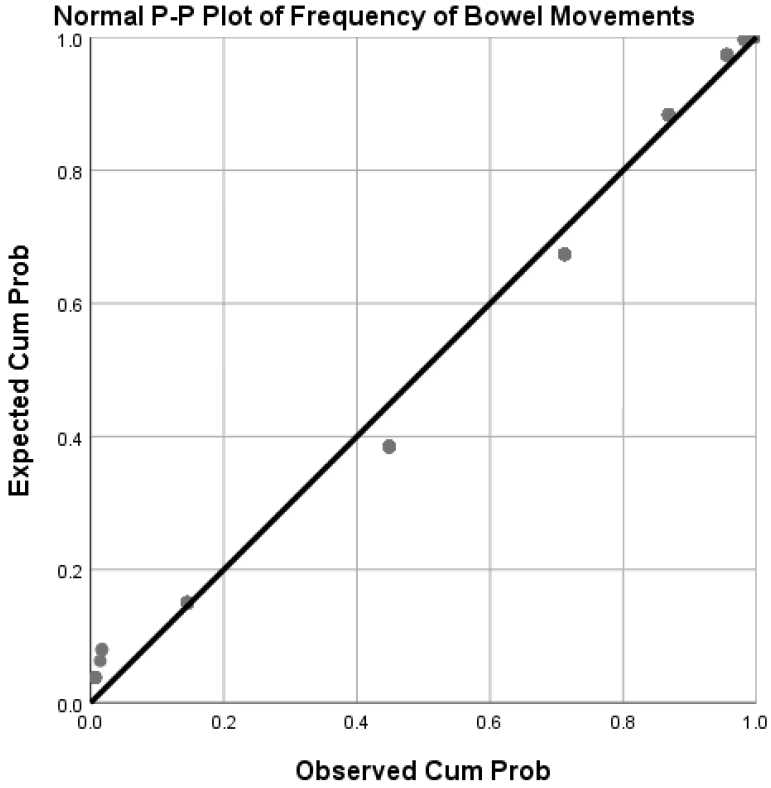
Histogram of Body Mass Index



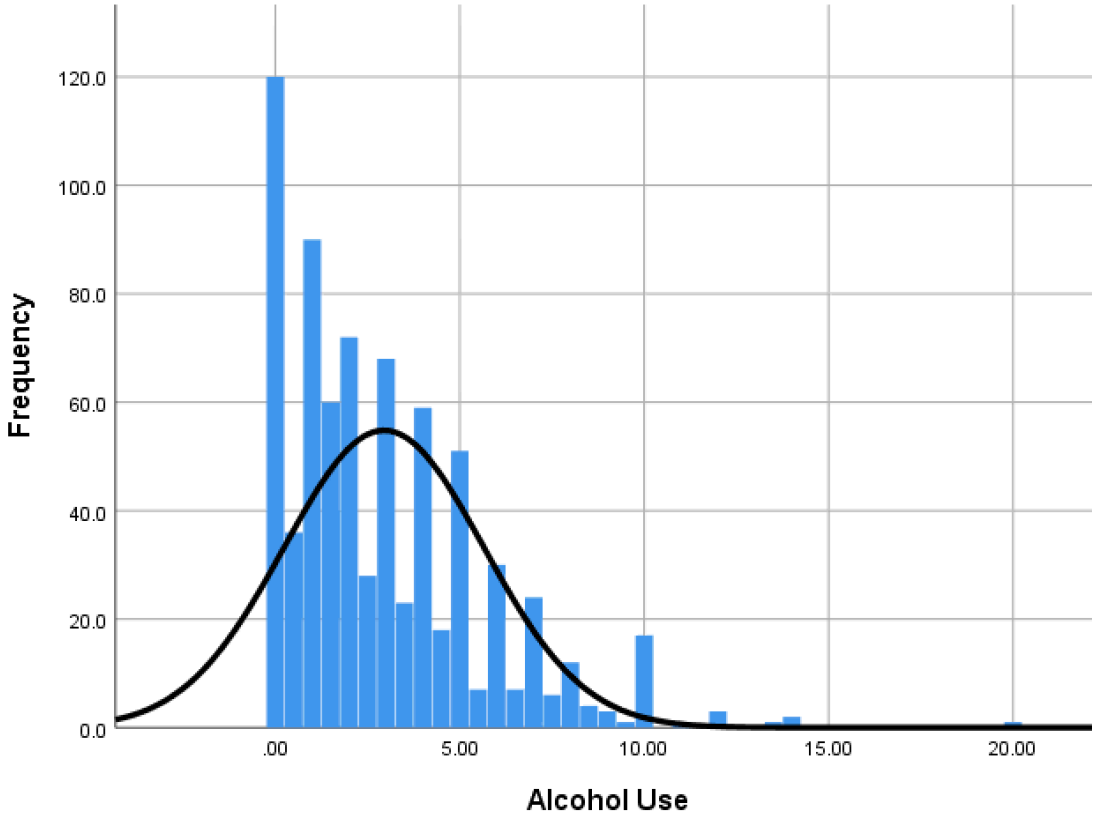


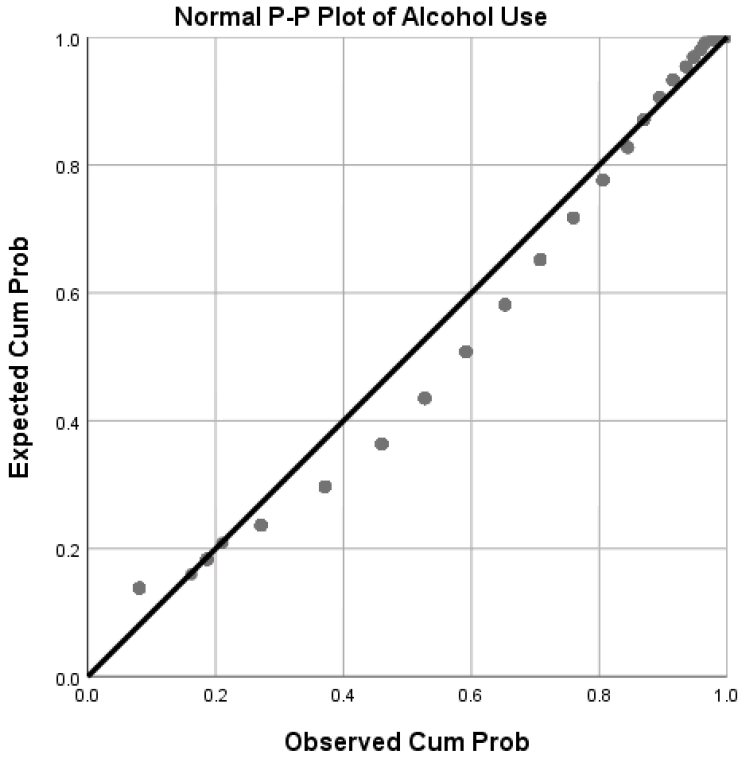
Histogram of Frequency of Bowel Movements



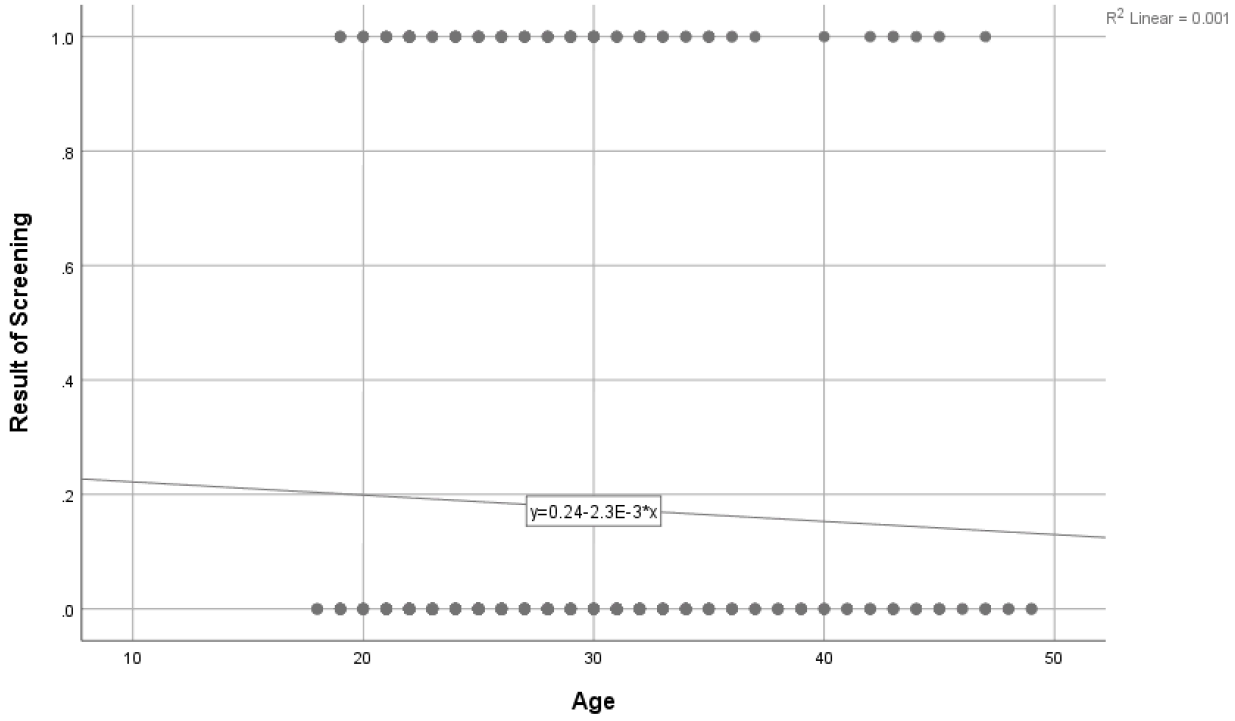


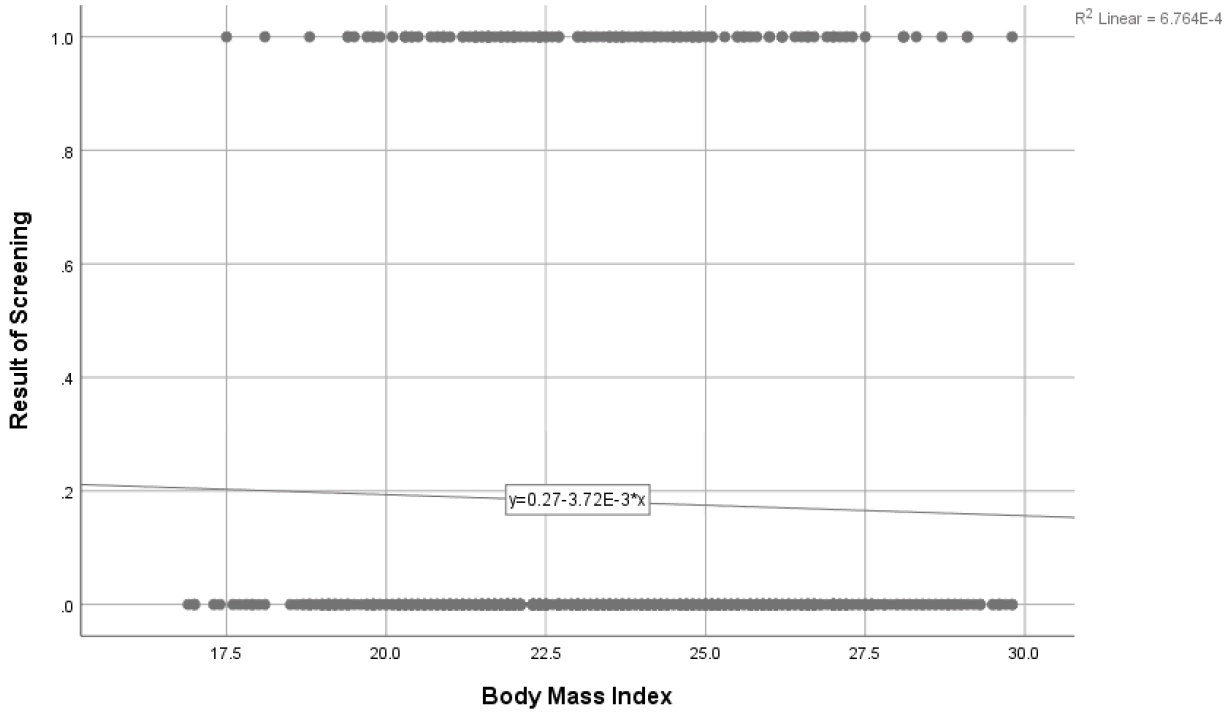
Histogram of Alcohol Use

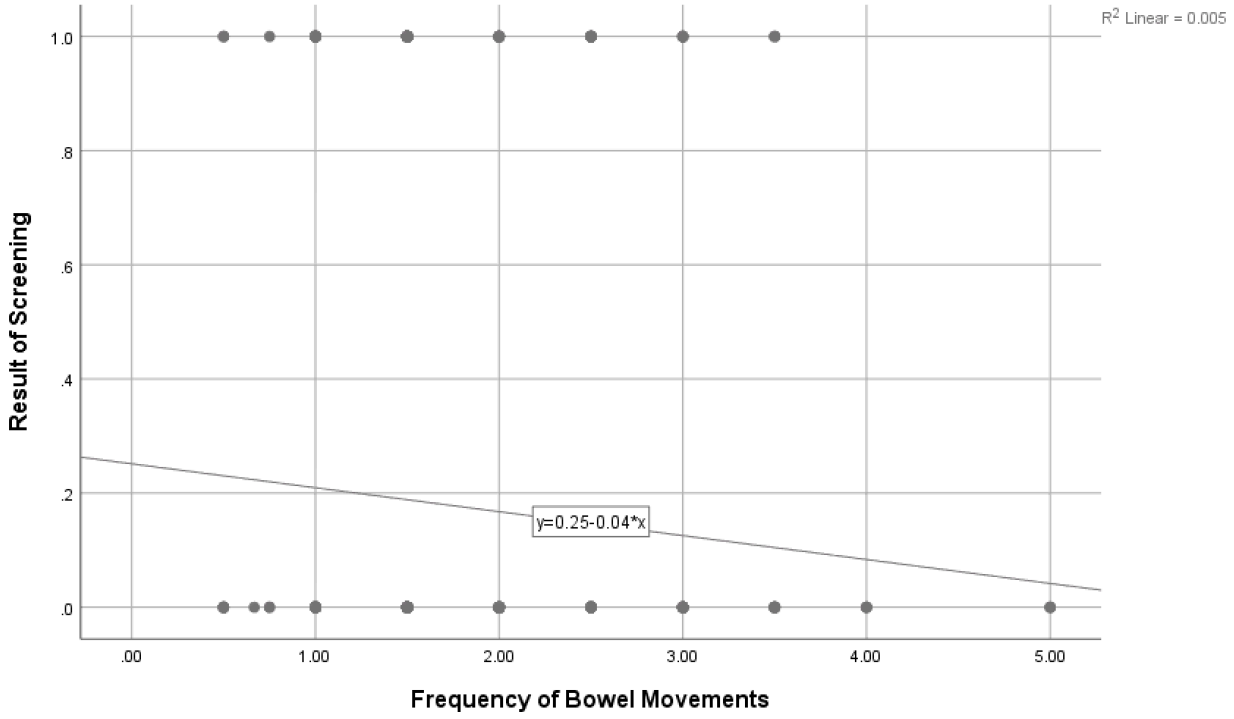


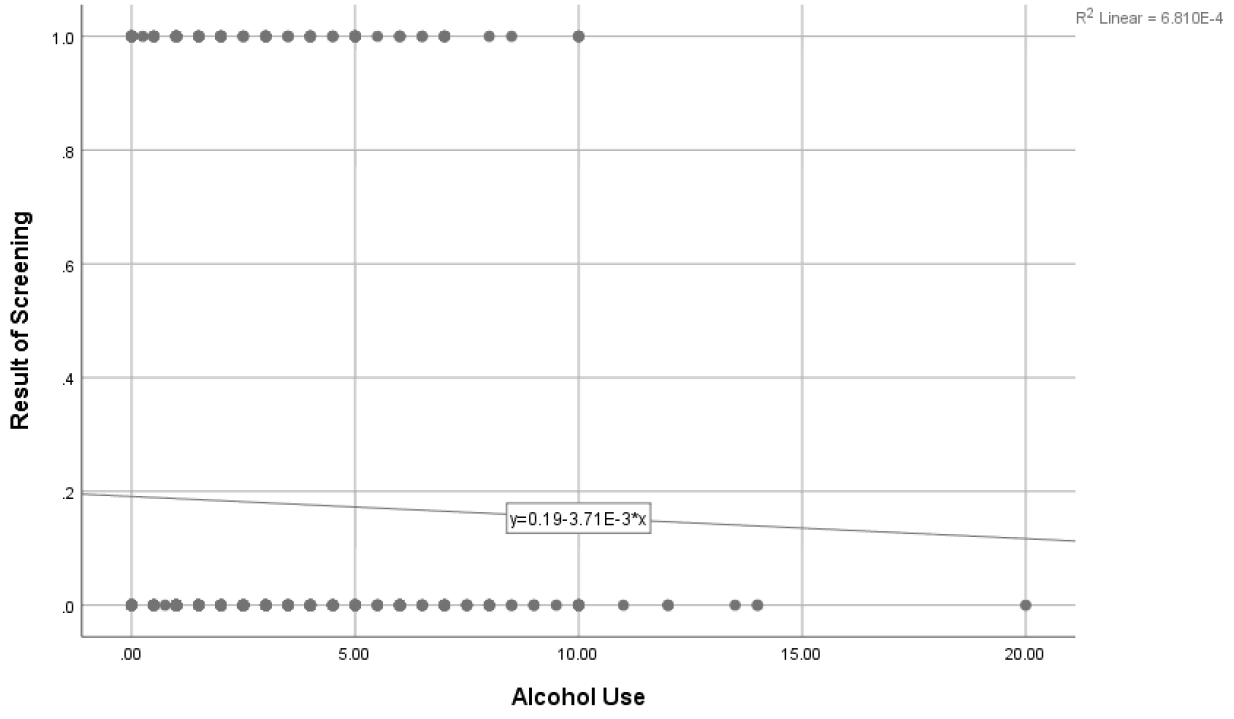


Appendix E. Scatterplots for continuous variables (age, body mass index, frequency of bowel movements, and alcohol use)

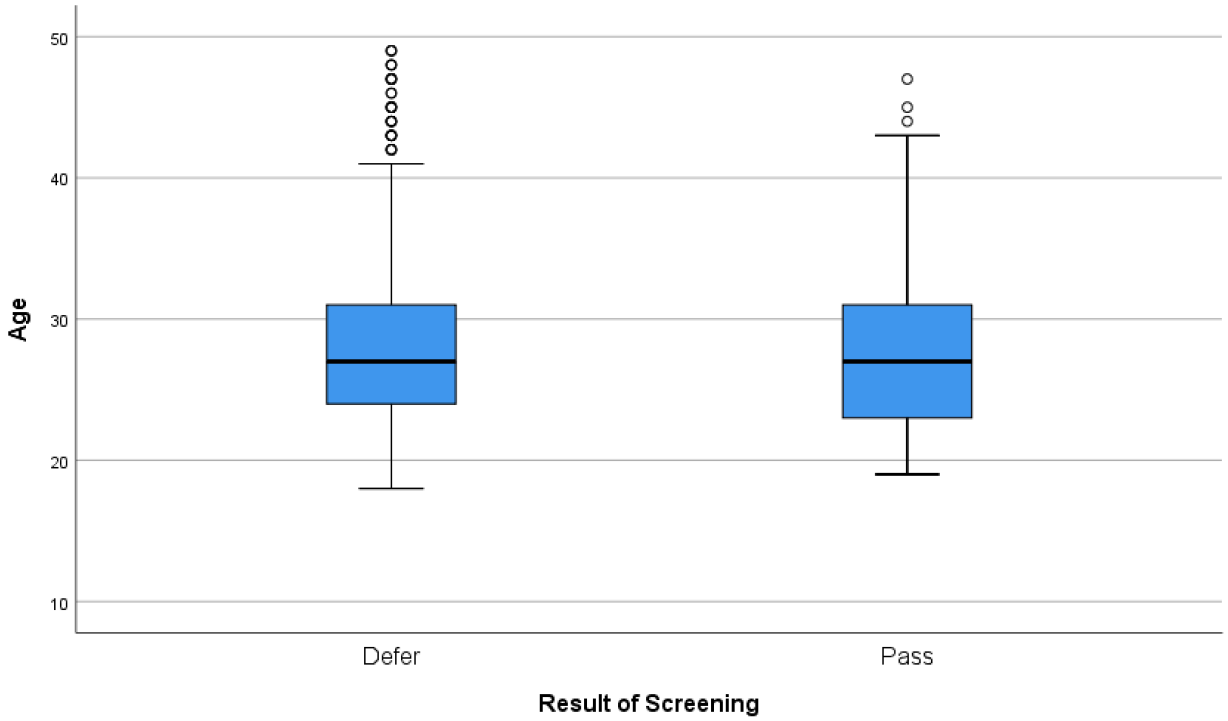


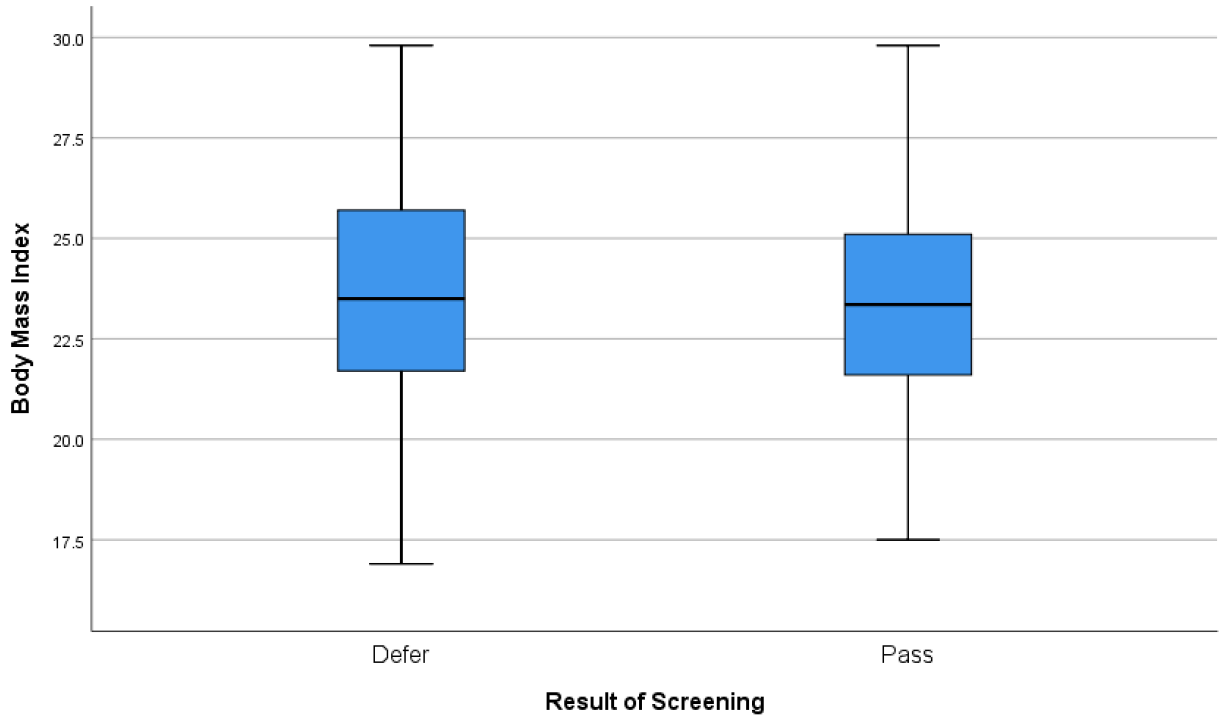


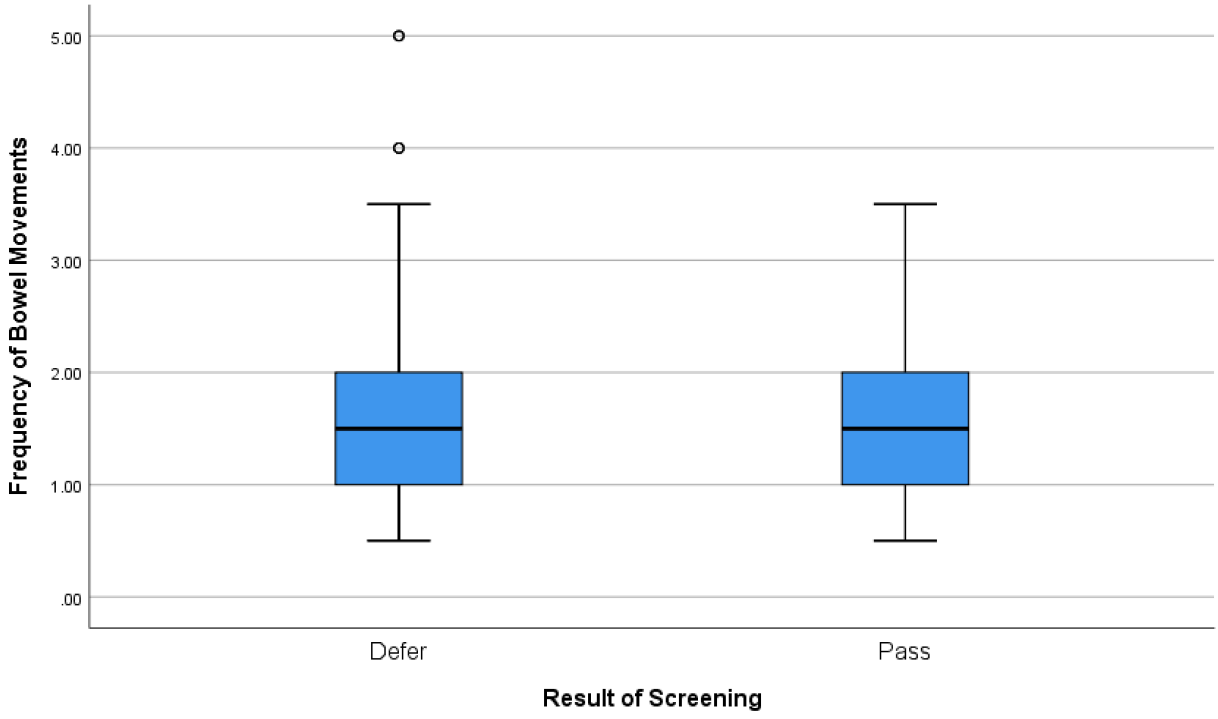

R² Linear = 0.005
Result of Screening
1.0
.8
.6
.4
.2
.0
y=0.25-0.04*x
.00
1.00
2.00
3.00
4.00
5.00
Frequency of Bowel Movements

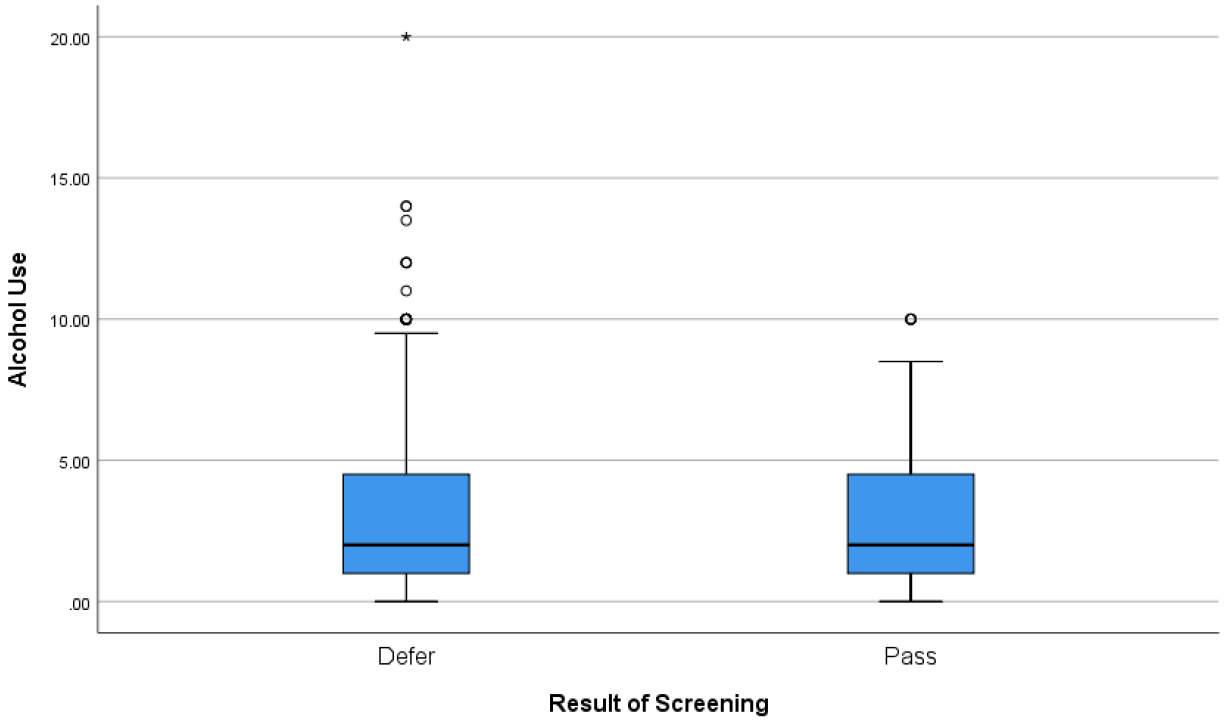

R² Linear = 6.810E-4
Result of Screening
1.0
.8
.6
.4
.2
.0
y=0.19-3.71E-3*x
.00
5.00
10.00
15.00
20.00
Alcohol Use

Appendix F. Boxplots for continuous variables (age, body mass index, frequency of bowel movements, and alcohol use) and results of clinical interview screening

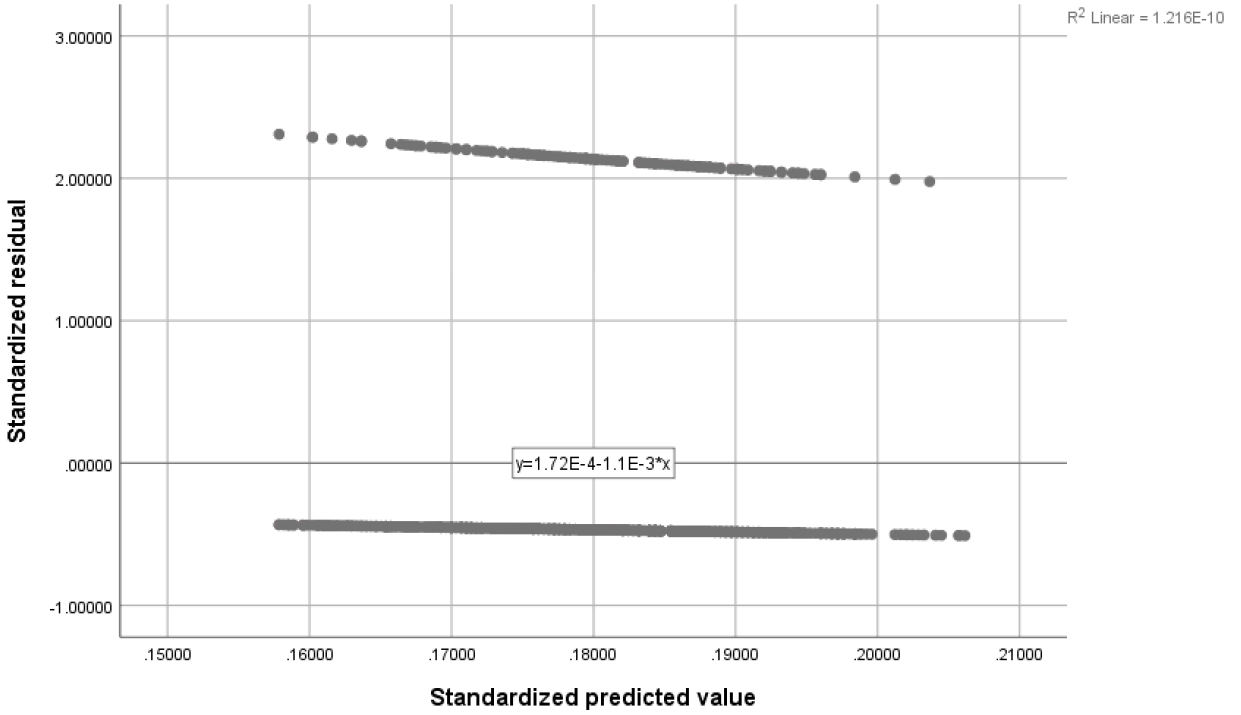


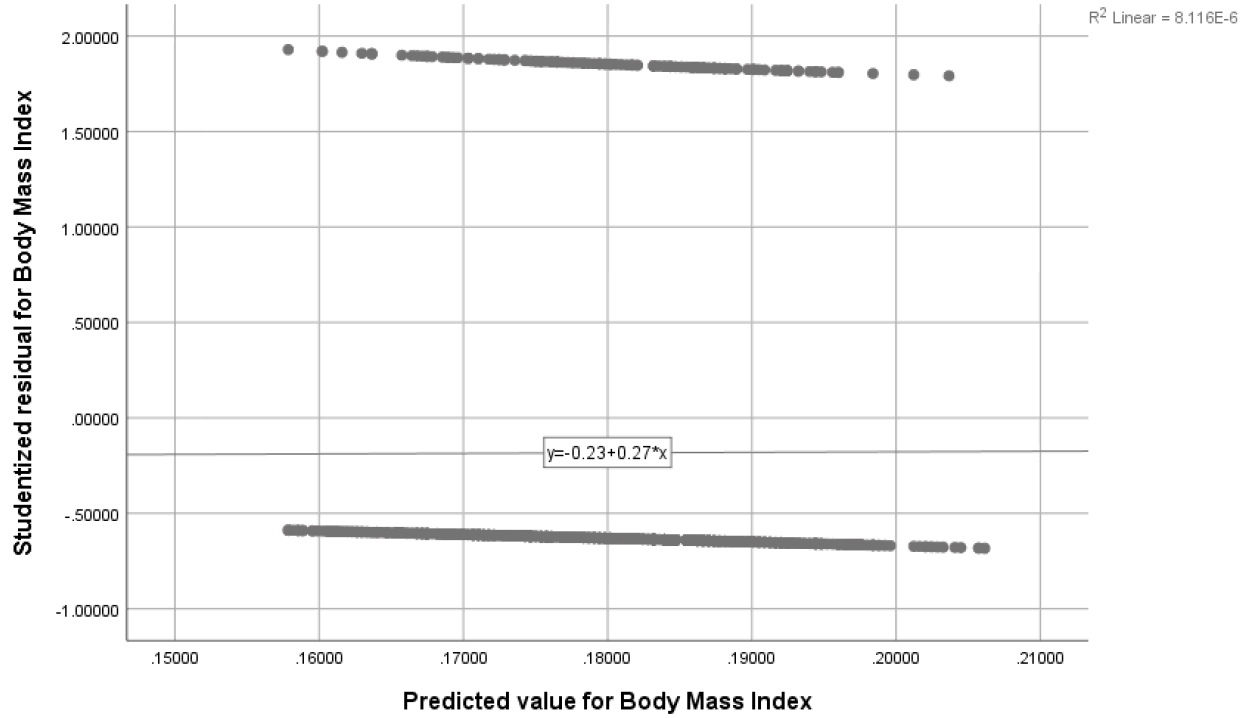


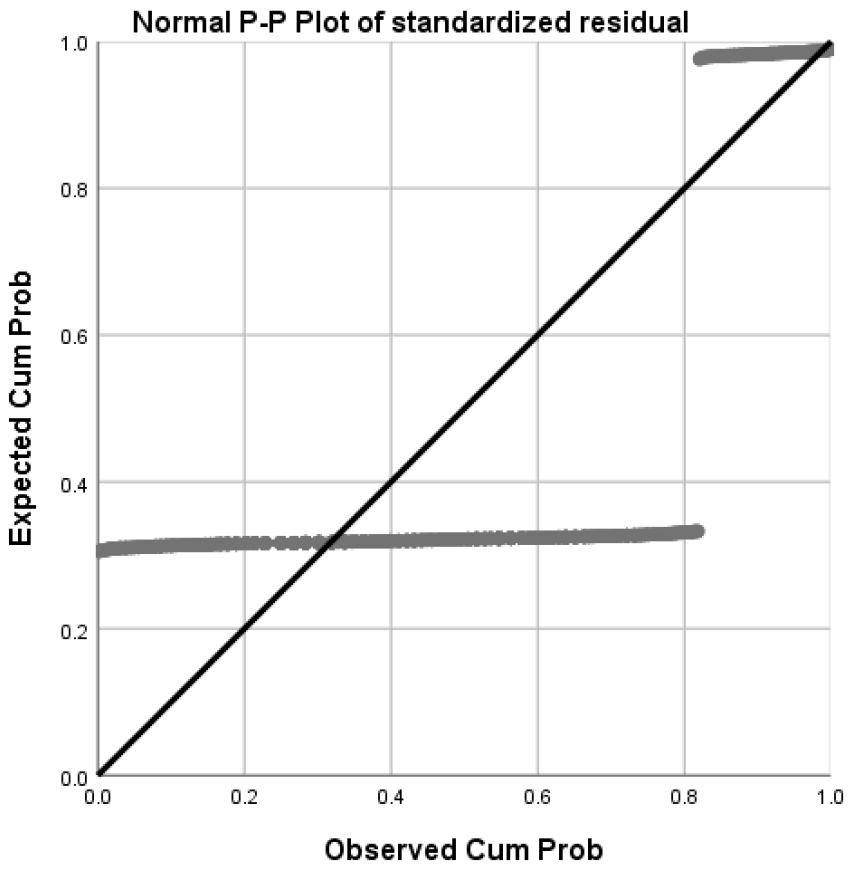

Frequency of Bowel Movements
5.00
4.00
3.00
2.00
1.00
.00
Defer
Pass
Result of Screening


Alcohol Use
20.00
15.00
10.00
5.00
.00
Defer
Pass
Result of Screening

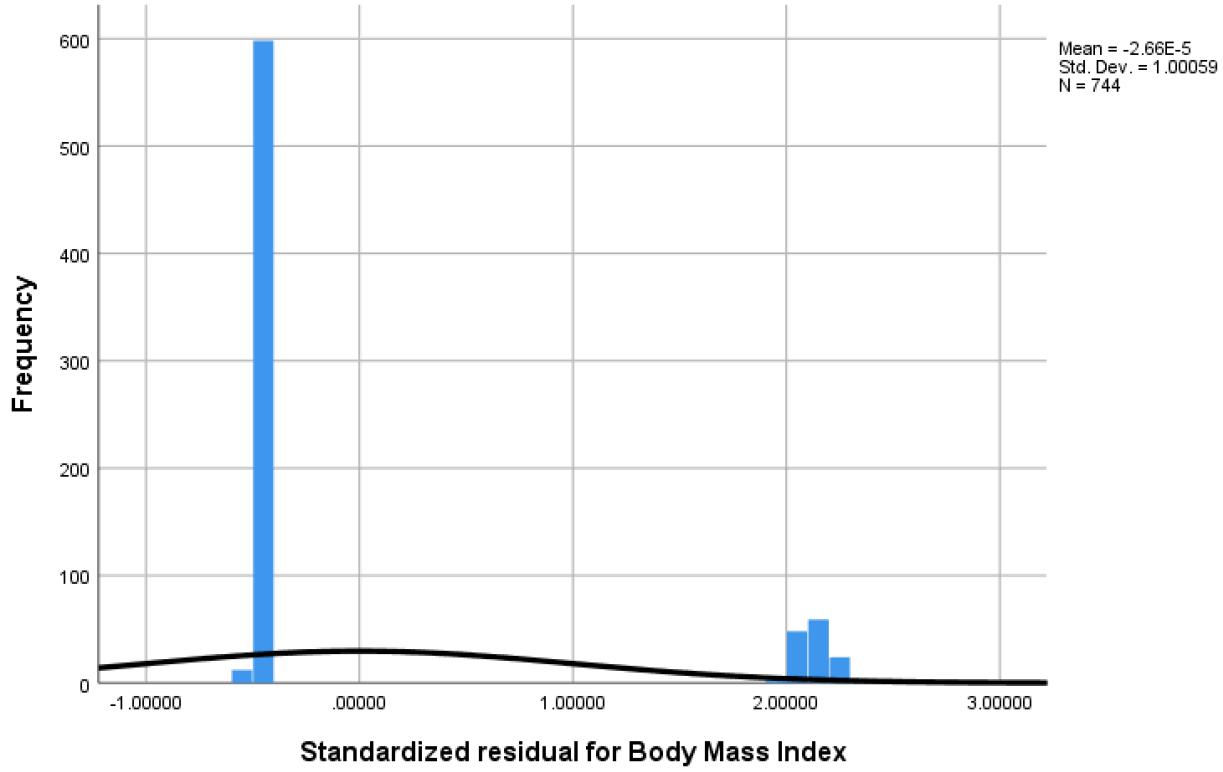
Appendix G. Scatterplot of standardized residuals against predicted values, body mass index






Normal P-P Plot of standardized residual
Expected Cum Prob
Observed Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0

Appendix H. Histogram of standardized residual for body mass index



Appendix I. Results from initial model for Aim 1 backward elimination logistic regression

Outcome variable coding

Original Value	Internal Value
Defer	0
Pass	1

Categorical variables codings

		Frequency	Parameter coding		
			-1	-2	-3
Seasonality	Winter	217	0	0	0
	Spring	139	1	0	0
	Summer	171	0	1	0
	Fall	217	0	0	1
Age	18 to 24 years	226	0	0	0
	25 to 29 years	272	1	0	0
	30 to 34 years	146	0	1	0
	35 to 50 years	100	0	0	1
Tobacco Use	Never	605	0	0	
	Former	87	1	0	
	Current	52	0	1	
Alcohol Use	None	120	0	0	
	Moderate	549	1	0	
	High	75	0	1	
Frequency of Bowel Movements	Fewer than two	465	0		
	Two of more	279	1		
Diet	Meat Based	645	0		
	No Meat	99	1		
Gender	Male	507	0		
	Female	237	1		

Block 0: Beginning Block

Iteration history[a,b,c]

Iteration		-2 Log likelihood	Coefficients Constant
Step 0	1	708.108	-1.280
	2	701.703	-1.500
	3	701.676	-1.516
	4	701.676	-1.516

a Constant is included in the model
b Initial -2 Log Likelihood: 701.676
c Estimation terminated at iteration number 4 because parameter estimates changed by less than 0.001

Classification table[a,b]

Observed			Predicted		
			Result of Screening		Percentage Correct
			Defer	Pass	
Step 0	Result of Screening	Defer	610	0	100
		Pass	134	0	0
	Overall Percentage				82

a Constant is included in the model
b The cut value is 0.500

Variables in the equation

		B	S.E.	Wald	df	Sig.	Exp (B)
Step 0	Constant	-1.516	0.095	252.372	1	0.000	0.220

Variables not in the equation

			Score	df	Sig.
Step 0	Variables	Gender (1)	0.921	1	0.337
		Age	1.524	3	0.677
		Age (1)	0.351	1	0.554
		Age (2)	0.168	1	0.682
		Age (3)	0.709	1	0.400
		Body Mass Index	0.503	1	0.478
		Frequency of BM (1)	5.828	1	0.016
		Diet (1)	1.841	1	0.175
		Alcohol Use	0.229	2	0.892
		Alcohol Use (1)	0.059	1	0.808
		Alcohol Use (2)	0.228	1	0.633
		Tobacco Use	1.128	2	0.569
		Tobacco Use (1)	0.246	1	0.620
		Tobacco Use (2)	0.784	1	0.376
		Seasonality	4.954	3	0.175
		Seasonality (1)	2.965	1	0.085
		Seasonality (2)	0.742	1	0.389
		Seasonality (3)	2.108	1	0.147
	Overall Statistics		17.501	14	0.230

Block 1: Method = Backward Stepwise (Likelihood Ratio)

Iteration history[a,b,c,d,e]

Iteration		-2 Log likelihood	Coefficients Constant
Step 1	1	694.403	-0.205
	2	683.756	0.223
	3	683.550	0.338
	4	683.550	0.341
	5	683.550	0.341
Step 2	1	694.749	-0.221
	2	684.225	0.188
	3	684.027	0.296
	4	684.026	0.299
	5	684.026	0.299
Step 3	1	695.841	-0.215
	2	685.611	0.210
	3	685.424	0.323
	4	685.423	0.326
	5	685.423	0.326
Step 4	1	696.728	-0.224
	2	686.801	0.199
	3	686.631	0.311
	4	686.631	0.314
	5	686.631	0.314
Step 5	1	697.482	-0.736
	2	687.800	-0.634
	3	687.642	-0.588
	4	687.641	-0.586
	5	687.641	-0.586
Step 6	1	700.711	-0.776
	2	692.009	-0.691
	3	691.894	-0.649
	4	691.894	-0.648
	5	691.894	-0.648
Step 7	1	702.075	-0.863
	2	693.769	-0.828
	3	693.668	-0.794
	4	693.668	-0.793
	5	693.668	-0.793

Iteration history[a,b,c,d,e] (continued)

Iteration		-2 Log likelihood	Coefficients Constant
Step 8	1	703.537	-0.893
	2	695.736	-0.876
	3	695.660	-0.846
	4	695.660	-0.846

a Method: Backward Stepwise (Likelihood Ratio)
b Constant is included in the model. Model too large to include all coefficient values
c Initial -2 Log Likelihood: 701.676
d Estimation terminated at iteration number 5 because parameter estimates changed by less than 0.001
e Estimation terminated at iteration number 4 because parameter estimates changed by less than 0.001

Omnibus tests of model coefficients

		Chi-square	df	Sig.
Step 1	Step	18.126	14	0.201
	Block	18.126	14	0.201
	Model	18.126	14	0.201
Step 2[a]	Step	-0.477	2	0.788
	Block	17.650	12	0.127
	Model	17.650	11	0.090
Step 3[a]	Step	-1.397	3	0.706
	Block	16.253	9	0.062
	Model	16.253	9	0.062
Step 4[a]	Step	-1.207	2	0.547
	Block	15.045	7	0.035
	Model	15.045	6	0.020
Step 5[a]	Step	-1.011	1	0.315
	Block	14.035	6	0.029
	Model	14.035	4	0.007
Step 6[a]	Step	-4.252	3	0.235
	Block	9.782	3	0.021
	Model	9.782	3	0.021
Step 7[a]	Step	-1.774	1	0.183
	Block	8.008	2	0.018
	Model	8.008	2	0.018
Step 8[a]	Step	-1.992	1	0.158
	Block	6.017	1	0.014
	Model	6.017	1	0.014

Model summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	683.550[a]	0.024	0.039
2	684.026[a]	0.023	0.038
3	685.423[a]	0.022	0.035
4	686.631[a]	0.020	0.033
5	687.641[a]	0.019	0.031
6	691.894[a]	0.013	0.021
7	693.668[a]	0.011	0.018
8	695.660[b]	0.008	0.013

a Estimation terminated at iteration number 5 because parameter estimates changed by less than 0.001

b Estimation terminated at iteration number 4 because parameter estimates changed by less than 0.001

Hosmer and Lemeshow test

Step	Chi-square	df	Sig.
1	8.234	8	0.411
2	2.52	8	0.961
3	7.597	8	0.474
4	4.47	8	0.812
5	6.118	8	0.634
6	0.808	4	0.937
7	1.244	2	0.537
8	0.00	0	.

Contingency table for Hosmer and Lemeshow test

		Result of Screening = Defer		Result of Screening = Pass		Total
		Observed	Expected	Observed	Expected	
Step 1	1	67	67	7	7	74
	2	67	65	7	9	74
	3	68	64	6	10	74
	4	58	63	16	11	74
	5	62	62	12	12	74
	6	60	60	14	14	74
	7	60	60	15	15	75
	8	59	57	15	17	74
	9	50	56	24	18	74
	10	59	55	18	22	77

Contingency table for Hosmer and Lemeshow test (continued)

		Result of Screening = Defer		Result of Screening = Pass		Total
		Observed	Expected	Observed	Expected	
Step 2	1	68	67	6	7	74
	2	66	65	8	9	74
	3	67	65	8	10	75
	4	61	63	13	11	74
	5	62	62	12	12	74
	6	58	60	16	14	74
	7	61	60	14	15	75
	8	58	58	17	17	75
	9	53	56	21	18	74
	10	56	53	19	22	75
Step 3	1	67	68	8	7	75
	2	68	65	6	9	74
	3	66	64	8	10	74
	4	59	63	15	11	74
	5	65	61	9	13	74
	6	60	60	14	14	74
	7	58	59	16	15	74
	8	58	58	16	16	74
	9	50	56	24	18	74
	10	59	56	18	21	77
Step 4	1	66	67	8	7	74
	2	67	65	7	9	74
	3	63	64	11	10	74
	4	62	63	12	11	74
	5	63	62	11	12	74
	6	64	60	10	14	74
	7	57	58	16	15	73
	8	55	58	20	17	75
	9	53	56	21	18	74
	10	60	57	18	21	78

Contingency table for Hosmer and Lemeshow test (continued)

		Result of Screening = Defer		Result of Screening = Pass		Total
		Observed	Expected	Observed	Expected	
Step 5	1	47	48	6	5	53
	2	70	67	6	9	76
	3	65	68	14	11	79
	4	65	64	11	12	76
	5	69	66	11	14	80
	6	65	61	9	13	74
	7	43	46	15	12	58
	8	68	69	19	18	87
	9	56	57	20	19	76
	10	62	63	23	22	85
Step 6	1	76	77	10	9	86
	2	23	23	4	4	27
	3	165	164	28	29	193
	4	28	29	7	6	35
	5	125	122	25	28	150
	6	193	194	60	59	253
Step 7	1	35	33	2	4	37
	2	206	208	36	34	242
	3	51	53	11	9	62
	4	318	316	85	87	403
Step 8	1	241	241	38	38	279
	2	369	369	96	96	465

Classification table[ab]

Observed			Predicted		
			Result of Screening		Percentage Correct
			Defer	Pass	
Step 1	Result of Screening	Defer	610	0	100
		Pass	134	0	0
	Overall Percentage				82
Step 2	Result of Screening	Defer	610	0	100
		Pass	134	0	0
	Overall Percentage				82
Step 3	Result of Screening	Defer	610	0	100
		Pass	134	0	0
	Overall Percentage				82
Step 4	Result of Screening	Defer	610	0	100
		Pass	134	0	0
	Overall Percentage				82
Step 5	Result of Screening	Defer	610	0	100
		Pass	134	0	0
	Overall Percentage				82
Step 6	Result of Screening	Defer	610	0	100
		Pass	134	0	0
	Overall Percentage				82
Step 7	Result of Screening	Defer	610	0	100
		Pass	134	0	0
	Overall Percentage				82
Step 8	Result of Screening	Defer	610	0	100
		Pass	134	0	0
	Overall Percentage				82

a Constant is included in the model
b The cut value is 0.500

Variables in the equation

		B	S.E.	Wald	df	Sig.	Exp (B)	95% C.I. for Exp(B)	
								Lower	Upper
Step 1[a]	Gender (1)	-0.341	0.224	2.311	1	0.128	0.711	0.459	1.104
	Age			1.477	3	0.688			
	Age (1)	-0.251	0.24	1.095	1	0.295	0.778	0.487	1.245
	Age (2)	-0.115	0.276	0.173	1	0.678	0.892	0.519	1.533
	Age (3)	-0.322	0.341	0.895	1	0.344	0.724	0.371	1.413
	Body Mass Index	-0.028	0.039	0.509	1	0.475	0.973	0.902	1.049
	Frequency of BM (1)	-0.509	0.215	5.592	1	0.018	0.601	0.394	0.917
	Diet (1)	-0.461	0.322	2.054	1	0.152	0.631	0.336	1.185
	Alcohol Use			0.459	2	0.795			
	Alcohol Use (1)	-0.027	0.267	0.01	1	0.921	0.974	0.577	1.645
	Alcohol Use (2)	-0.249	0.405	0.378	1	0.539	0.78	0.353	1.724
	Tobacco Use			1.121	2	0.571			
	Tobacco Use (1)	-0.11	0.319	0.119	1	0.730	0.896	0.48	1.673
	Tobacco Use (2)	-0.438	0.428	1.048	1	0.306	0.645	0.279	1.493
	Seasonality			4.244	3	0.236			
	Seasonality (1)	-0.491	0.309	2.528	1	0.112	0.612	0.334	1.121
	Seasonality (2)	-0.297	0.277	1.148	1	0.284	0.743	0.432	1.279
	Seasonality (3)	0.045	0.243	0.035	1	0.852	1.047	0.65	1.685
	Constant	0.341	0.997	0.117	1	0.732	1.406		
Step 2[a]	Gender (1)	-0.326	0.223	2.139	1	0.144	0.722	0.466	1.117
	Age			1.397	3	0.706			
	Age (1)	-0.242	0.239	1.027	1	0.311	0.785	0.491	1.254
	Age (2)	-0.112	0.276	0.164	1	0.686	0.894	0.52	1.537
	Age (3)	-0.313	0.338	0.854	1	0.356	0.732	0.377	1.42
	Body Mass Index	-0.028	0.039	0.534	1	0.465	0.972	0.901	1.049
	Frequency of BM (1)	-0.507	0.215	5.56	1	0.018	0.602	0.395	0.918
	Diet (1)	-0.447	0.32	1.948	1	0.163	0.64	0.342	1.198
	Tobacco Use			1.238	2	0.538			
	Tobacco Use (1)	-0.127	0.317	0.16	1	0.690	0.881	0.473	1.64
	Tobacco Use (2)	-0.454	0.427	1.13	1	0.288	0.635	0.275	1.467
	Seasonality			4.208	3	0.240			
	Seasonality (1)	-0.485	0.309	2.474	1	0.116	0.615	0.336	1.127
	Seasonality (2)	-0.294	0.277	1.127	1	0.288	0.745	0.433	1.283
	Seasonality (3)	0.049	0.243	0.041	1	0.839	1.051	0.652	1.692
	Constant	0.299	0.967	0.095	1	0.758	1.348		

Variables in the equation (continued)

		B	S.E.	Wald	df	Sig.	Exp (B)	95% C.I. for Exp(B)	
								Lower	Upper
Step 3[a]	Gender (1)	-0.333	0.222	2.249	1	0.134	0.717	0.464	1.108
	Body Mass Index	-0.036	0.038	0.928	1	0.335	0.964	0.896	1.038
	Frequency of BM (1)	-0.498	0.214	5.401	1	0.020	0.608	0.4	0.925
	Diet (1)	-0.444	0.319	1.935	1	0.164	0.641	0.343	1.199
	Tobacco Use			1.132	2	0.568			
	Tobacco Use (1)	-0.159	0.315	0.257	1	0.612	0.853	0.46	1.58
	Tobacco Use (2)	-0.413	0.425	0.945	1	0.331	0.662	0.288	1.521
	Seasonality			4.308	3	0.230			
	Seasonality (1)	-0.481	0.308	2.439	1	0.118	0.618	0.338	1.131
	Seasonality (2)	-0.294	0.276	1.131	1	0.287	0.745	0.434	1.281
	Seasonality (3)	0.06	0.242	0.062	1	0.803	1.062	0.661	1.708
	Constant	0.326	0.955	0.116	1	0.733	1.385		
Step 4[a]	Gender (1)	-0.321	0.222	2.1	1	0.147	0.725	0.47	1.12
	Body Mass Index	-0.038	0.038	1.007	1	0.316	0.963	0.894	1.037
	Frequency of BM (1)	-0.502	0.214	5.515	1	0.019	0.605	0.398	0.92
	Diet (1)	-0.441	0.319	1.909	1	0.167	0.644	0.345	1.203
	Seasonality			4.248	3	0.236			
	Seasonality (1)	-0.474	0.308	2.375	1	0.123	0.622	0.34	1.138
	Seasonality (2)	-0.283	0.275	1.059	1	0.303	0.754	0.44	1.291
	Seasonality (3)	0.068	0.242	0.08	1	0.777	1.071	0.667	1.719
	Constant	0.314	0.956	0.108	1	0.743	1.369		
Step 5[a]	Gender (1)	-0.268	0.215	1.554	1	0.213	0.765	0.502	1.166
	Frequency of BM (1)	-0.517	0.213	5.888	1	0.015	0.596	0.392	0.905
	Diet (1)	-0.41	0.317	1.671	1	0.196	0.664	0.356	1.236
	Seasonality			4.096	3	0.251			
	Seasonality (1)	-0.464	0.307	2.284	1	0.131	0.628	0.344	1.148
	Seasonality (2)	-0.255	0.273	0.87	1	0.351	0.775	0.454	1.324
	Seasonality (3)	0.082	0.241	0.114	1	0.735	1.085	0.677	1.74
	Constant	-0.586	0.334	3.083	1	0.079	0.556		
Step 6[a]	Gender (1)	-0.282	0.214	1.731	1	0.188	0.754	0.496	1.148
	Frequency of BM (1)	-0.548	0.212	6.69	1	0.010	0.578	0.381	0.876
	Diet (1)	-0.409	0.316	1.674	1	0.196	0.665	0.358	1.234
	Constant	-0.648	0.31	4.377	1	0.036	0.523		
Step 7[a]	Frequency of BM (1)	-0.502	0.209	5.773	1	0.016	0.605	0.402	0.912
	Diet (1)	-0.427	0.315	1.839	1	0.175	0.652	0.352	1.21
	Constant	-0.793	0.291	7.458	1	0.006	0.452		

Variables in the equation (continued)

		B	S.E.	Wald	df	Sig.	Exp (B)	95% C.I. for Exp(B)	
								Lower	Upper
Step 8[a]	Frequency of BM (1)	-0.501	0.209	5.753	1	0.016	0.606	0.403	0.913
	Constant	-0.846	0.288	8.62	1	0.003	0.429		

a Variable(s) entered on step 1: Gender, Age, Body Mass Index, Frequency of BM, Diet, Alcohol Use, Tobacco Use, Seasonality

Model if term removed

Variable		Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of the Change
Step 1	Gender	-342.96	2.37	1	0.124
	Age	-342.514	1.477	3	0.688
	Body Mass Index	-342.03	0.51	1	0.475
	Frequency of BM (1)	-344.686	5.823	1	0.016
	Diet	-342.888	2.226	1	0.136
	Alcohol Use	-342.013	0.477	2	0.788
	Tobacco Use	-342.379	1.209	2	0.546
	Seasonality	-343.978	4.407	3	0.221
Step 2	Gender	-343.11	2.193	1	0.139
	Age	-342.712	1.397	3	0.706
	Body Mass Index	-342.281	0.535	1	0.465
	Frequency of BM (1)	-344.908	5.79	1	0.016
	Diet	-343.067	2.108	1	0.147
	Tobacco Use	-342.682	1.337	2	0.513
	Seasonality	-344.197	4.368	3	0.224
Step 3	Gender	-343.865	2.307	1	0.129
	Body Mass Index	-343.178	0.932	1	0.334
	Frequency of BM (1)	-345.524	5.624	1	0.018
	Diet	-343.759	2.094	1	0.148
	Tobacco Use	-343.315	1.207	2	0.547
	Seasonality	-344.946	4.468	3	0.215
Step 4	Gender	-344.392	2.152	1	0.142
	Body Mass Index	-343.821	1.011	1	0.315
	Frequency of BM (1)	-346.189	5.746	1	0.017
	Diet	-344.348	2.065	1	0.151
	Seasonality	-345.518	4.405	3	0.221
Step 5	Gender	-344.615	1.589	1	0.207
	Frequency of BM (1)	-346.894	6.147	1	0.013
	Diet	-344.721	1.8	1	0.180
	Seasonality	-345.947	4.252	3	0.235

Model if term removed (continued)

Variable		Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of the Change
Step 6	Gender	-346.834	1.774	1	0.183
	Frequency of BM (1)	-349.452	7.01	1	0.008
	Diet	-346.85	1.805	1	0.179
Step 7	Frequency of BM (1)	-349.853	6.038	1	0.014
	Diet	-347.83	1.992	1	0.158
Step 8	Frequency of BM (1)	-350.838	6.017	1	0.014

Variables not in the equation

			Score	df	Sig.
Step 2[a]	Variables	Alcohol Use	0.461	2	0.794
		Alcohol Use (1)	0.093	1	0.761
		Alcohol Use (2)	0.451	1	0.502
	Overall Statistics		0.461	2	0.794
Step 3[b]	Variables	Age	1.401	1	0.705
		Age (2)	0.044	1	0.834
		Age (3)	0.353	1	0.552
		Alcohol Use	0.384	2	0.825
		Alcohol Use (1)	0.120	1	0.729
		Alcohol Use (2)	0.384	1	0.536
	Overall Statistics		1.861	5	0.868
Step 4[c]	Variables	Age	1.271	3	0.736
		Age (1)	0.483	1	0.487
		Age (2)	0.047	1	0.828
		Age (3)	0.321	1	0.571
		Alcohol Use	0.487	2	0.784
		Alcohol Use (1)	0.113	1	0.736
		Alcohol Use (2)	0.482	1	0.487
		Tobacco Use	1.142	2	0.565
		Tobacco Use (1)	0.178	1	0.673
		Tobacco Use (2)	0.882	1	0.348
	Overall Statistics		3.002	7	0.885

Variables not in the equation (continued)

			Score	df	Sig.
Step 5[d]	Variables	Age	1.658	3	0.646
		Age (1)	0.497	1	0.481
		Age (2)	0.039	1	0.844
		Age (3)	0.528	1	0.467
		Body Mass Index	1.009	1	0.315
		Alcohol Use	0.516	2	0.773
		Alcohol Use (1)	0.157	1	0.692
		Alcohol Use (2)	0.516	1	0.473
		Tobacco Use	1.216	2	0.545
		Tobacco Use (1)	0.193	1	0.661
		Tobacco Use (2)	0.935	1	0.334
	Overall Statistics		4.011	8	0.856
Step 6[e]	Variables	Age	1.740	3	0.628
		Age (1)	0.635	1	0.425
		Age (2)	0.058	1	0.809
		Age (3)	0.461	1	0.497
		Body Mass Index	0.857	1	0.355
		Alcohol Use	0.498	2	0.780
		Alcohol Use (1)	0.185	1	0.667
		Alcohol Use (2)	0.496	1	0.481
		Tobacco Use	1.157	2	0.561
		Tobacco Use (1)	0.202	1	0.653
		Tobacco Use (2)	0.862	1	0.353
		Seasonality	4.141	3	0.247
		Seasonality (1)	2.453	1	0.117
		Seasonality (2)	0.638	1	0.424
		Seasonality (3)	1.789	1	0.181
	Overall Statistics		8.121	11	0.702

Variables not in the equation (continued)

			Score	df	Sig.
Step 7[f]	Variables	Gender (1)	1.738	1	0.187
		Age	1.735	3	0.629
		Age (1)	0.513	1	0.474
		Age (2)	0.055	1	0.815
		Age (3)	0.575	1	0.448
		Body Mass Index	0.345	1	0.557
		Alcohol Use	0.300	2	0.861
		Alcohol Use (1)	0.101	1	0.751
		Alcohol Use (2)	0.300	1	0.584
		Tobacco Use	1.004	2	0.605
		Tobacco Use (1)	0.201	1	0.654
		Tobacco Use (2)	0.719	1	0.396
		Seasonality	4.324	3	0.229
		Seasonality (1)	2.550	1	0.110
		Seasonality (2)	0.635	1	0.425
		Seasonality (3)	2.022	1	0.155
	Overall Statistics		9.852	12	0.629
Step 8[g]	Variables	Gender (1)	1.918	1	0.166
		Age	1.679	3	0.642
		Age (1)	0.422	1	0.516
		Age (2)	0.052	1	0.820
		Age (3)	0.638	1	0.424
		Body Mass Index	0.197	1	0.657
		Diet (1)	1.862	1	0.172
		Alcohol Use	0.226	2	0.893
		Alcohol Use (1)	0.124	1	0.725
		Alcohol Use (2)	0.212	1	0.645
		Tobacco Use	0.975	2	0.614
		Tobacco Use (1)	0.240	1	0.624
		Tobacco Use (2)	0.648	1	0.421
		Seasonality	4.354	3	0.226
		Seasonality (1)	2.580	1	0.108
		Seasonality (2)	0.576	1	0.448
		Seasonality (3)	2.198	1	0.138
	Overall Statistics		11.703	13	0.552

a Variable(s) removed on step 2: Alcohol Use
b Variable(s) removed on step 3: Age
c Variable(s) removed on step 4: Tobacco Use
d Variable(s) removed on step 5: Body Mass Index
e Variable(s) removed on step 6: Seasonality
f Variable(s) removed on step 7: Gender
g Variable(s) removed on step 8: Diet

Casewise list[b]

Case	Selected Status[a]	Observed	Predicted	Predicted Group	Temporary Variable		
		Result of Screening			Resid	ZResid	SResid
20	S	P**	0.136	D	0.864	2.518	2
21	S	P**	0.136	D	0.864	2.518	2
42	S	P**	0.136	D	0.864	2.518	2
43	S	P**	0.136	D	0.864	2.518	2
46	S	P**	0.136	D	0.864	2.518	2
59	S	P**	0.136	D	0.864	2.518	2
61	S	P**	0.136	D	0.864	2.518	2
71	S	P**	0.136	D	0.864	2.518	2
85	S	P**	0.136	D	0.864	2.518	2
94	S	P**	0.136	D	0.864	2.518	2
97	S	P**	0.136	D	0.864	2.518	2
98	S	P**	0.136	D	0.864	2.518	2
103	S	P**	0.136	D	0.864	2.518	2
115	S	P**	0.136	D	0.864	2.518	2
119	S	P**	0.136	D	0.864	2.518	2
121	S	P**	0.136	D	0.864	2.518	2
419	S	P**	0.136	D	0.864	2.518	2
421	S	P**	0.136	D	0.864	2.518	2
444	S	P**	0.136	D	0.864	2.518	2
480	S	P**	0.136	D	0.864	2.518	2
519	S	P**	0.136	D	0.864	2.518	2
520	S	P**	0.136	D	0.864	2.518	2
548	S	P**	0.136	D	0.864	2.518	2
590	S	P**	0.136	D	0.864	2.518	2
592	S	P**	0.136	D	0.864	2.518	2
593	S	P**	0.136	D	0.864	2.518	2
595	S	P**	0.136	D	0.864	2.518	2
619	S	P**	0.136	D	0.864	2.518	2
624	S	P**	0.136	D	0.864	2.518	2
627	S	P**	0.136	D	0.864	2.518	2
628	S	P**	0.136	D	0.864	2.518	2
653	S	P**	0.136	D	0.864	2.518	2
685	S	P**	0.136	D	0.864	2.518	2
708	S	P**	0.136	D	0.864	2.518	2
713	S	P**	0.136	D	0.864	2.518	2
726	S	P**	0.136	D	0.864	2.518	2
728	S	P**	0.136	D	0.864	2.518	2

Casewise list[b] (continued)

Case	Selected Status[a]	Observed	Predicted	Predicted Group	Temporary Variable		
		Result of Screening			Resid	ZResid	SResid
731	S	P**	0.136	D	0.864	2.518	2

a S = Selected, U = Unselected cases, and ** = Misclassified cases
b Cases with studentized residuals greater than 2.000 are listed

Case Summaries for first 15 cases

Case Number	Donor ID	Result of Screening Coded	Frequency of BM (1)	Predicted probability	Predicted group
1	1	Pass	Fewer than two	0.20645	Defer
2	4	Defer	Fewer than two	0.20645	Defer
3	5	Pass	Fewer than two	0.20645	Defer
4	11	Defer	Fewer than two	0.20645	Defer
5	13	Defer	Two or more	0.13620	Defer
6	14	Pass	Fewer than two	0.20645	Defer
7	18	Defer	Fewer than two	0.20645	Defer
8	19	Defer	Fewer than two	0.20645	Defer
9	21	Defer	Fewer than two	0.20645	Defer
10	23	Defer	Two or more	0.13620	Defer
11	24	Defer	Two or more	0.13620	Defer
12	25	Pass	Fewer than two	0.20645	Defer
13	26	Pass	Fewer than two	0.20645	Defer
14	27	Defer	Two or more	0.13620	Defer
15	28	Pass	Fewer than two	0.20645	Defer

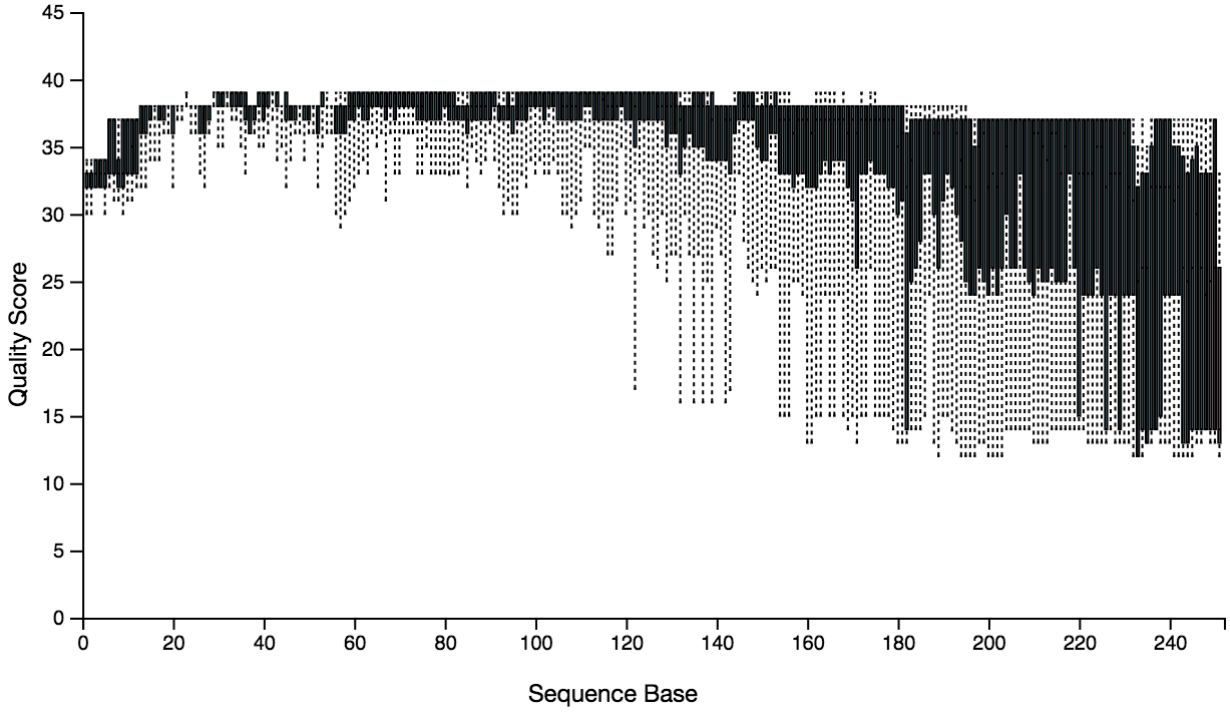
Appendix J. Results of forced entry simple logistic regression models

Model Variables (Step 1)	Beta	S.E.	Wald	*p*	Odds Ratio	CI 95%	
						Lower	Upper
Gender							
Female	0.201	0.210	0.919	0.338	1.223	0.810	1.847
Constant	-1.656	0.177	87.47	0.000*	0.191		
Age							
18 to 24 years of age			1.518	0.678			
25 to 29 years of age	0.343	0.326	1.107	0.293	1.409	0.744	2.668
30 to 34 years of age	0.143	0.323	0.195	0.659	1.153	0.612	2.174
35 to 50 years of age	0.296	0.350	0.715	0.398	1.345	0.677	2.671
Constant	-1.735	0.280	38.363	0.000*	0.176		
Body Mass Index	-0.025	0.036	0.503	0.478	0.975	0.909	1.046
Constant	-0.921	0.841	1.199	0.274	0.398		
Frequency of BM							
Two or more per day	0.501	0.209	5.753	0.016*	1.650	1.096	2.484
Constant	-1.847	0.175	112.003	0.000*	0.158		
Diet							
No Meat	-0.424	0.314	1.818	0.178	0.655	0.354	1.212
Constant	-1.042	0.360	8.401	0.004*	0.353		
Alcohol Use							
None			0.229	0.892			
Moderate	0.164	0.394	0.174	0.676	1.179	0.545	2.549
High	0.156	0.334	0.219	0.639	1.169	0.608	2.249
Constant	-1.658	0.315	27.717	0.000*	0.190		
Tobacco Use							
None			1.119	0.572			
Former	0.390	0.419	0.863	0.353	1.476	0.649	3.360
Current	0.209	0.500	0.175	0.676	1.233	0.463	3.286
Constant	-1.861	0.406	20.974	0.000*	0.156		
Seasonality							
Winter			4.891	0.180			
Spring	-0.085	0.238	0.127	0.721	0.919	0.576	1.464
Summer	-0.592	0.302	3.839	0.050	0.553	0.306	1.000
Fall	-0.361	0.268	1.820	0.177	0.697	0.413	1.177
Constant	-1.313	0.166	62.494	0.000*	0.269		

* Significant at the *p* value level of <0.05

Appendix K. QIIME2 generated quality plots

Forward Reads

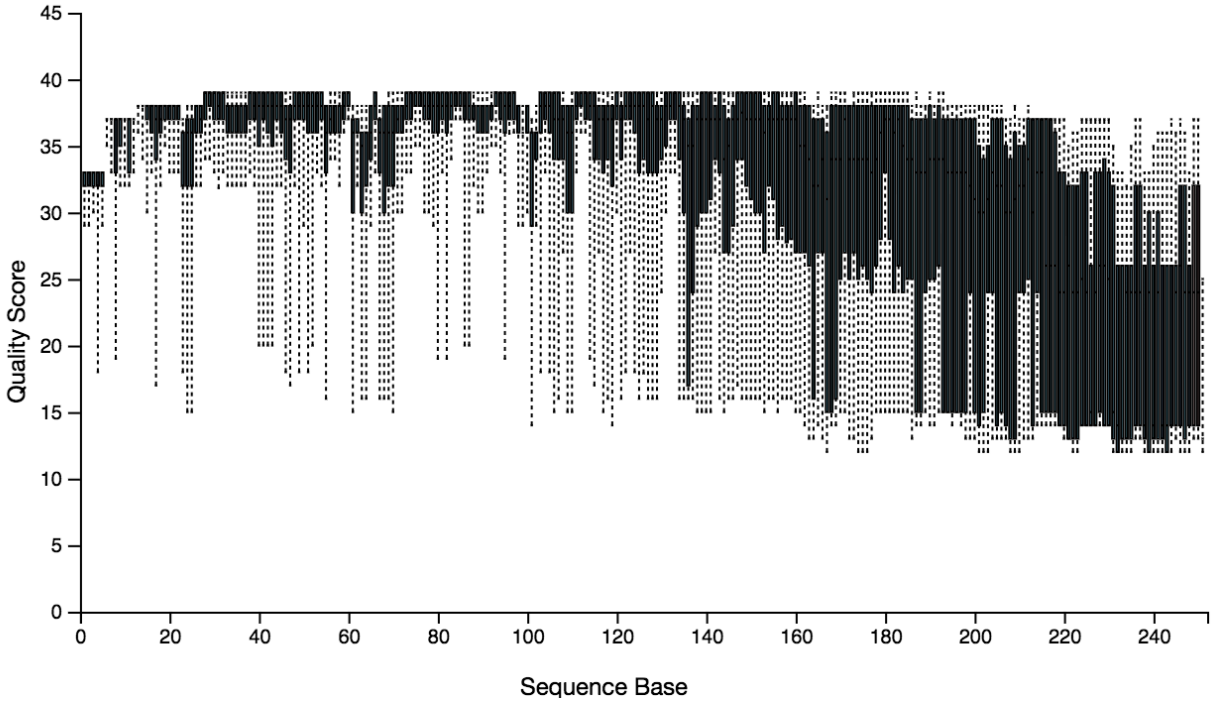


The plot at position 200 was generated using a random sampling of 10,000 out of 3,048,790 sequences without replacement. The minimum sequence length identified during subsampling was 248 bases. Outlier quality scores are not shown in box plots for clarity.

Parametric seven-number summary for position 200

Box Plot Feature	Percentile	Quality Score
(Not shown in box plot)	2nd	12
Lower Whisker	9th	12
Bottom of Box	25th	24
Middle of Box	50th (Median)	32
Top of Box	75th	37
Upper Whisker	91st	37
(Not shown in box plot)	98th	38

Reverse Reads



The plot at position 200 was generated using a random sampling of 10,000 out of 3,048,790 sequences without replacement. The minimum sequence length identified during subsampling was 248 bases. Outlier quality scores are not shown in box plots for clarity.

Parametric seven-number summary for position 200

Box Plot Feature	Percentile	Quality Score
(Not shown in box plot)	2nd	12
Lower Whisker	9th	13
Bottom of Box	25th	15
Middle of Box	50th (Median)	31
Top of Box	75th	37
Upper Whisker	91st	38
(Not shown in box plot)	98th	39

Appendix L. Change in descriptive statistics for continuous variables between Aim 1 and Aim 2

		N	Range	Minimum	Maximum
Age	Aim 1	744	31.0	18.0	49.0
	Aim 2	87	26.0	19.0	45.0
	Change		5.0	-1.0	4.0
Body Mass Index	Aim 1	744	12.9	16.9	29.8
	Aim 2	87	11.6	17.5	29.1
	Change		1.30	-0.6	0.7
Frequency of Bowel Movements	Aim 1	744	4.5	0.5	5.0
	Aim 2	87	3.0	0.5	3.5
	Change		1.5	0.0	1.5
Alcohol Use	Aim 1	744	20.0	0.0	20.0
	Aim 2	87	10.0	0.0	10.0
	Change		10.0	0.0	10.0

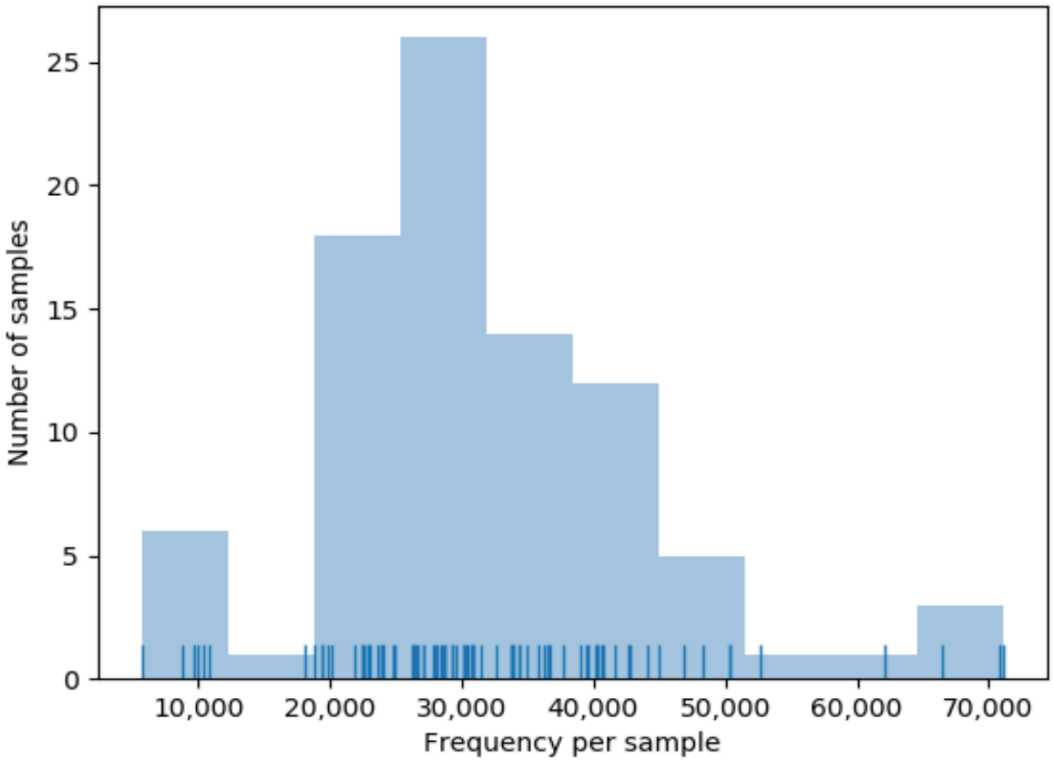
		Mean		Std. Deviation	Variance
		Statistic	Std. Error		
Age	Aim 1	28.1	0.23	6.19	38.25
	Aim 2	27.3	0.60	5.48	30.00
	Change	0.8	-0.40	0.70	8.25
Body Mass Index	Aim 1	23.6	0.10	2.69	7.22
	Aim 2	23.2	0.30	2.55	6.49
	Change	0.4	-0.17	0.14	0.73
Frequency of Bowel Movements	Aim 1	1.7	0.02	0.67	0.45
	Aim 2	1.6	0.10	0.67	0.45
	Change	0.1	0.00	0.00	0.00
Alcohol Use	Aim 1	2.9	0.10	2.71	7.33
	Aim 2	3.1	0.30	2.47	6.12
	Change	-0.1	-0.20	0.20	1.21

Appendix M. Change in descriptive statistics for categorical variables between Aim 1 and Aim 2

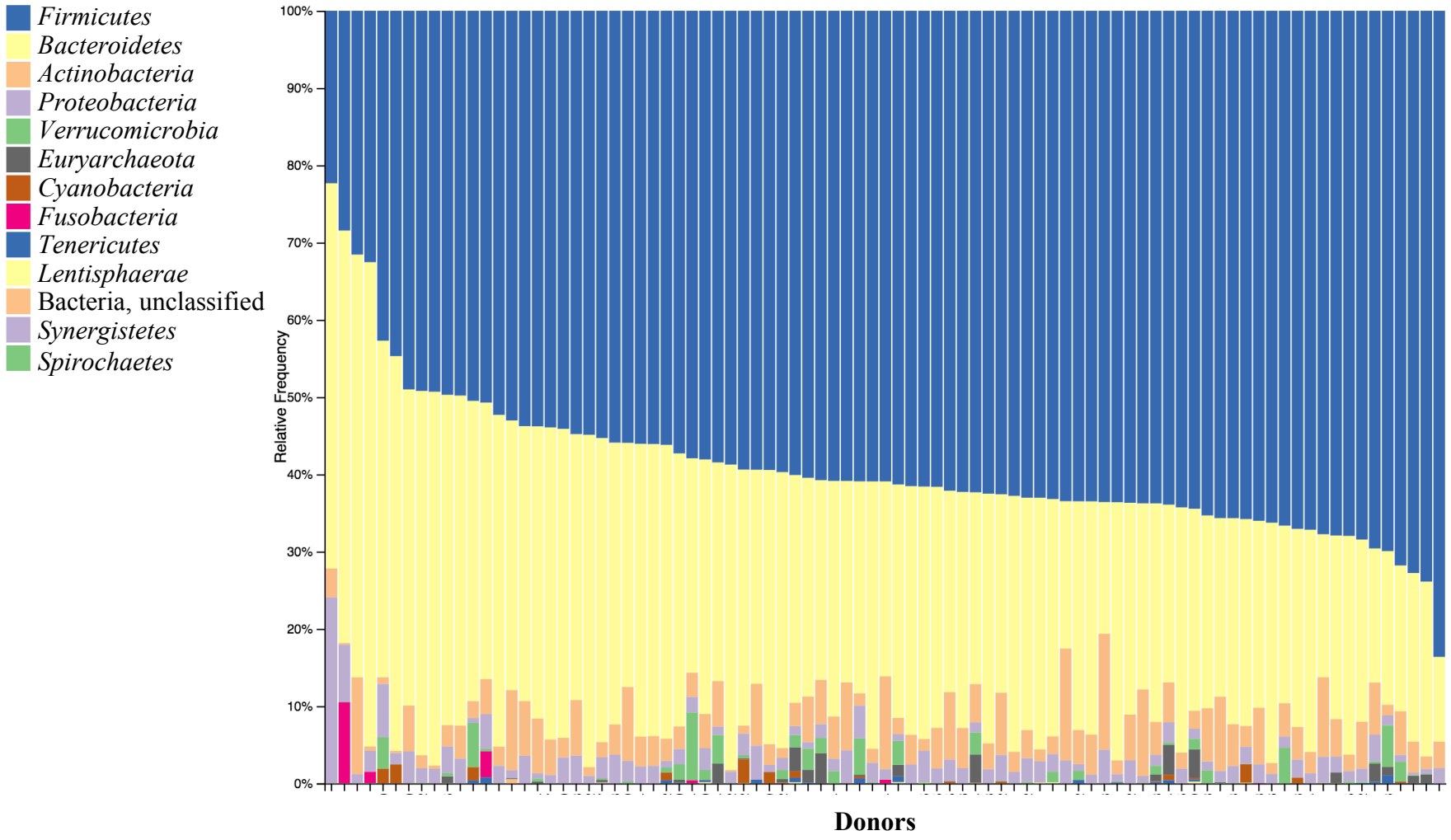
			Frequency	Percent	Cumulative Percent
Gender	Male	Aim 1	507	68.1%	68.1%
		Aim 2	63	72.4%	72.4%
		Change	444	-4.3%	
	Female	Aim 1	237	31.9%	100.0%
		Aim 2	24	27.6%	100.0%
		Change	213	4.3%	
	Total	Aim 1	744		
		Aim 2	87		
		Change	657	88.3%	
Diet	Meat Based	Aim 1	645	86.7%	86.7%
		Aim 2	79	90.8%	90.8%
		Change	566	-4.1%	
	No Meat	Aim 1	99	13.3%	100.0%
		Aim 2	8	9.2%	100.0%
		Change	91	4.1%	
	Total	Aim 1	744		
		Aim 2	87		
		Change	657	88.3%	
Tobacco Use	Never	Aim 1	605	81.3%	81.3%
		Aim 2	73	83.9%	83.9%
		Change	532	-2.6%	
	Former	Aim 1	87	11.7%	93.0%
		Aim 2	8	9.2%	93.1%
		Change	79	2.5%	
	Current	Aim 1	52	7.0%	100.0%
		Aim 2	6	6.9%	100.0%
		Change	46	0.1%	
	Subtotal	Aim 1	744		
		Aim 2	87		
		Change	657	88.3%	

			Frequency	Percent	Cumulative Percent
Seasonality	Winter	Aim 1	217	29.2%	29.2%
		Aim 2	29	33.3%	33.3%
		Change	188	-4.2%	
	Spring	Aim 1	139	18.7%	47.8%
		Aim 2	11	12.6%	46.0%
		Change	128	6.0%	
	Summer	Aim 1	171	23.0%	70.8%
		Aim 2	17	19.5%	65.5%
		Change	154	3.4%	
	Fall	Aim 1	217	29.2%	100.0%
		Aim 2	30	34.5%	100.0%
		Change	187	-5.3%	
	Total	Aim 1	744		
		Aim 2	87		
		Change	657	88.3%	

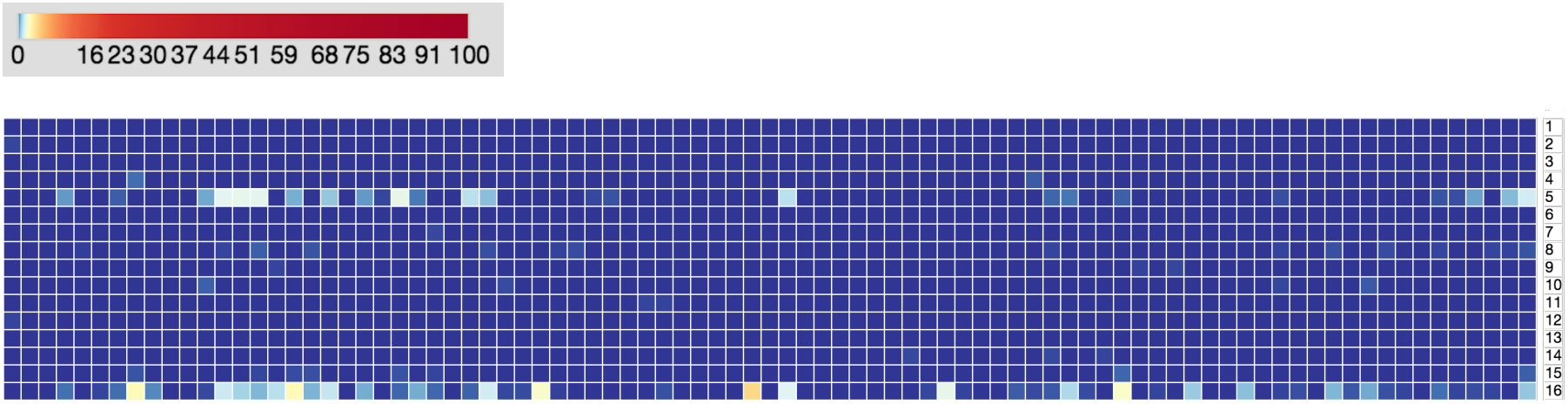
Appendix N. Histogram of bacterial operational taxonomy unit frequency by sample



Appendix O. Operational taxonomic unit bar chart by donor

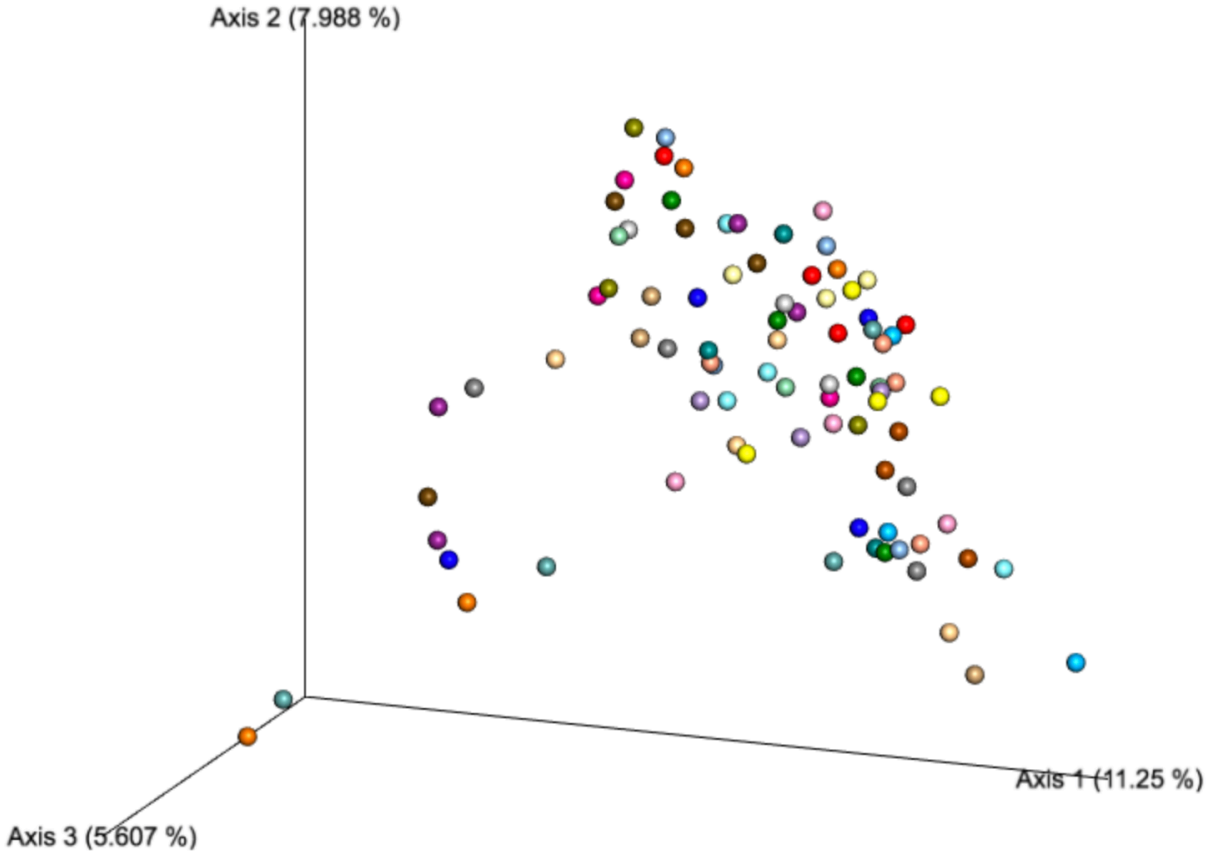


Appendix P. Heatmap of phylum levels

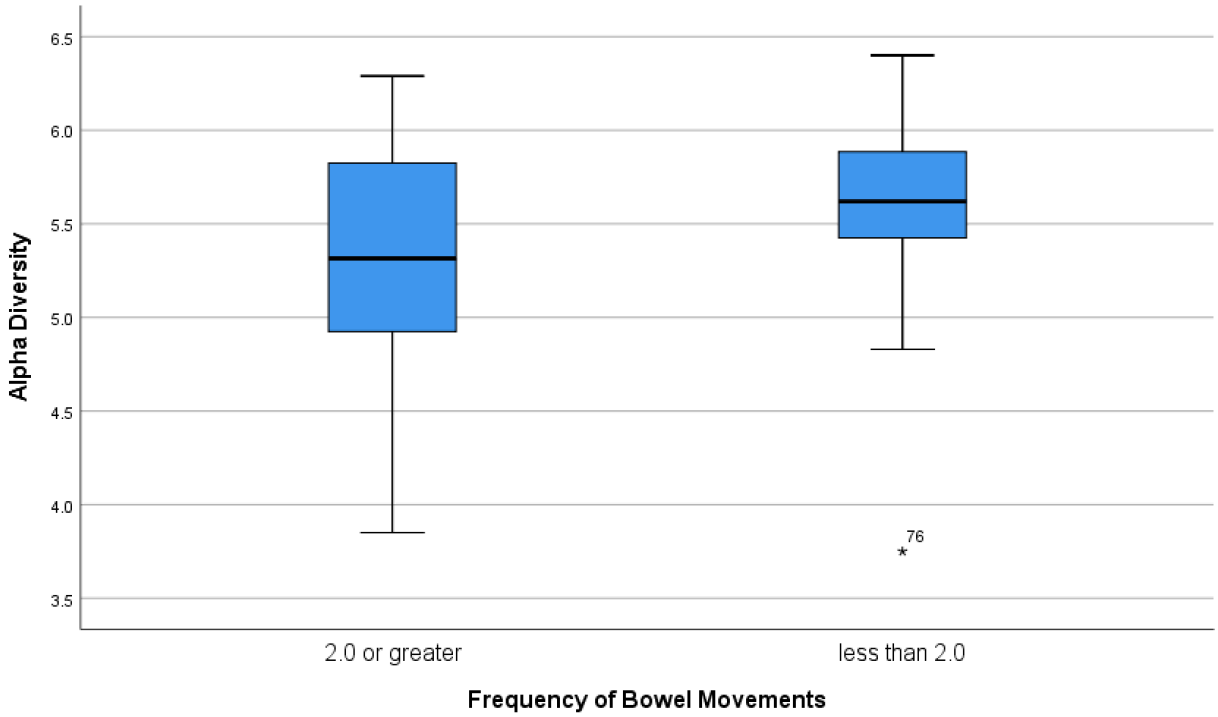


1. *Actinobacteria*
2. Bacteria
3. *Bacteroidetes*
4. *Cyanobacteria*
5. *Euryarchaeota*
6. *Firmicutes*
7. *Fusobacteria*
8. *Lentisphaerae*
9. Bacteria, unclassified OTU1492
10. Bacteria, unclassified OTU16
11. Bacteria, unclassified OTU443
12. Bacteria, unclassified OTU860
13. *Proteobacteria*
14. *Synergistetes*
15. *Tenericutes*
16. *Verrucomicrobia*

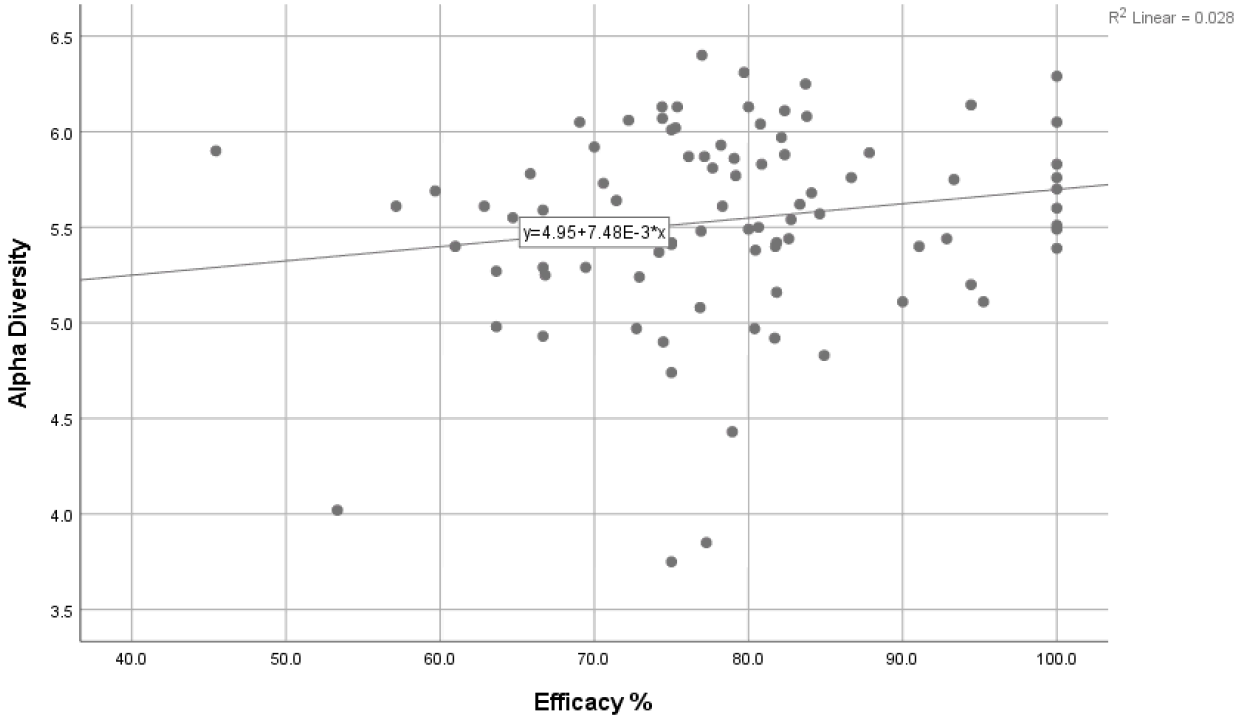
Appendix Q. Principal coordinates of analyses (PCoA) utilizing bray-curtis, emperor plot of donor 16S rRNA microbiome habitats



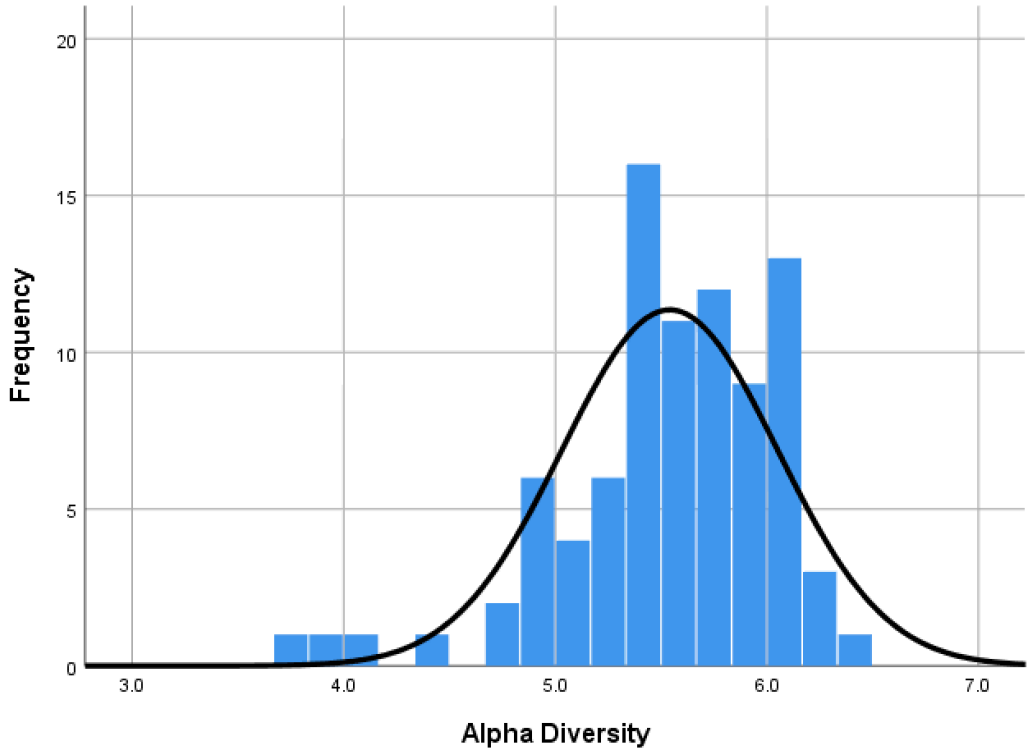
Appendix R. Box and whisker plot of alpha diversity per frequency of bowel movements per day category

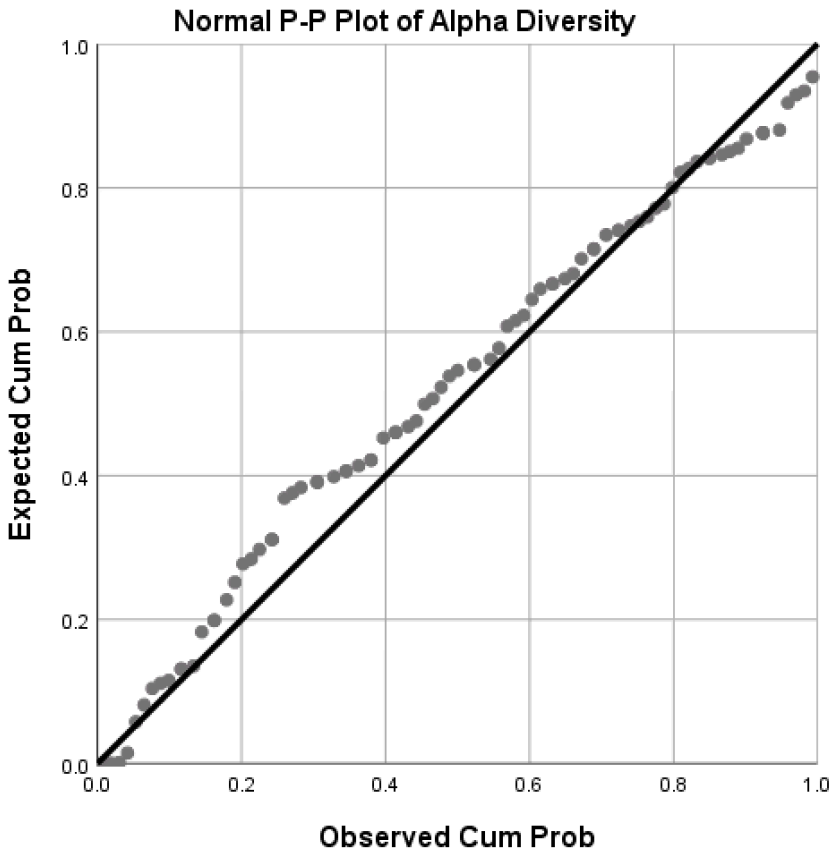


Appendix S. Simple scatterplot of alpha diversity and efficacy

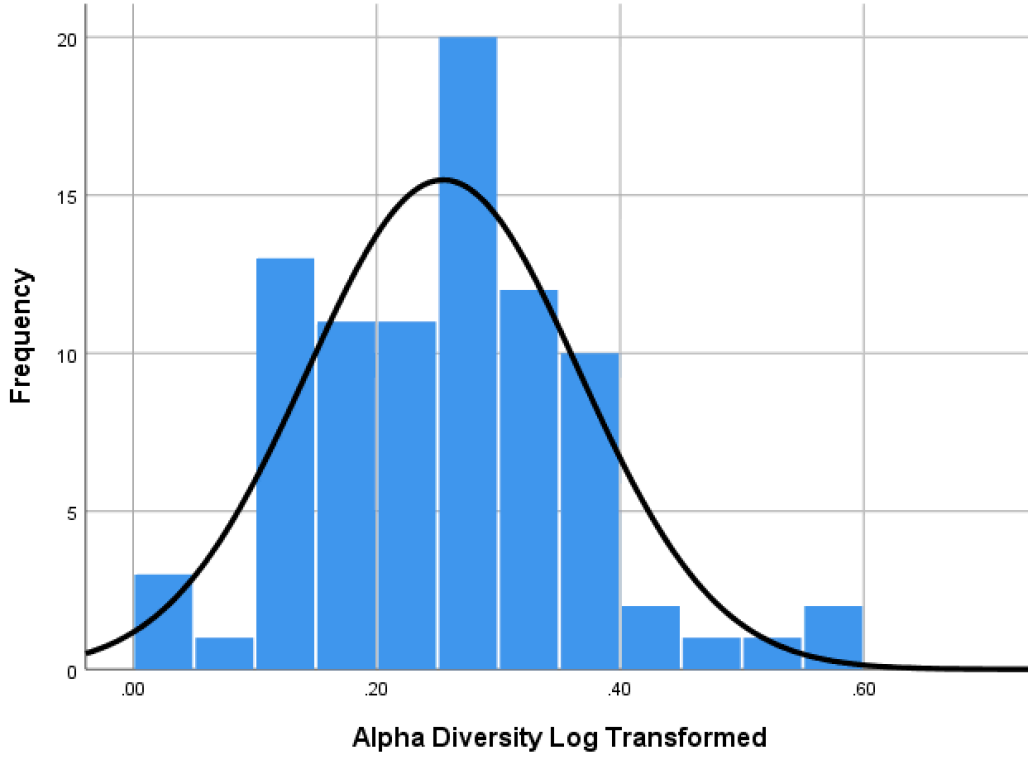


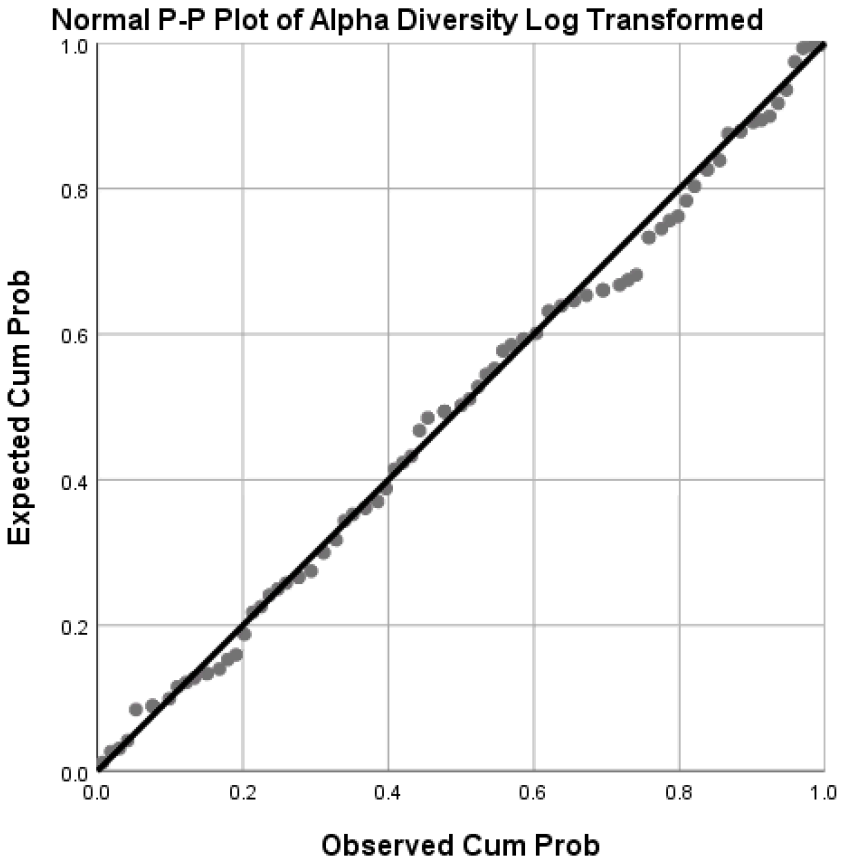
Appendix T. Histogram and P-P plot for alpha diversity



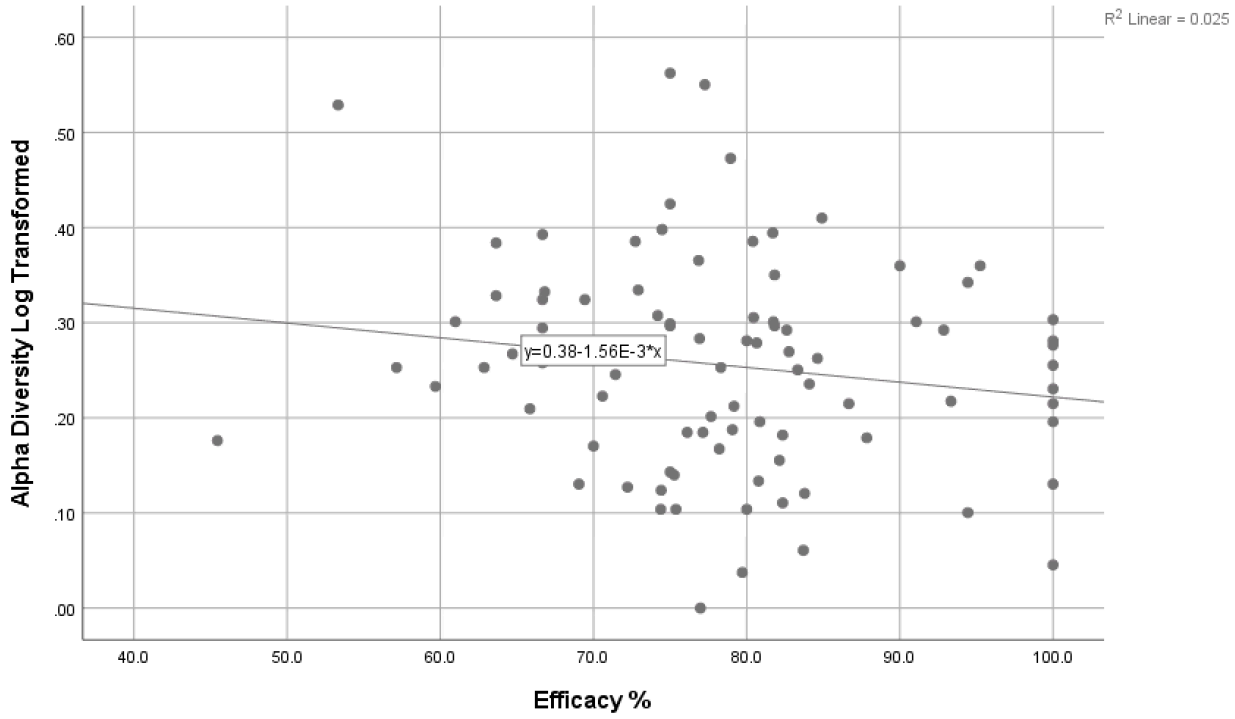


Appendix U. Histogram and P-P plot for alpha diversity, log-transformed

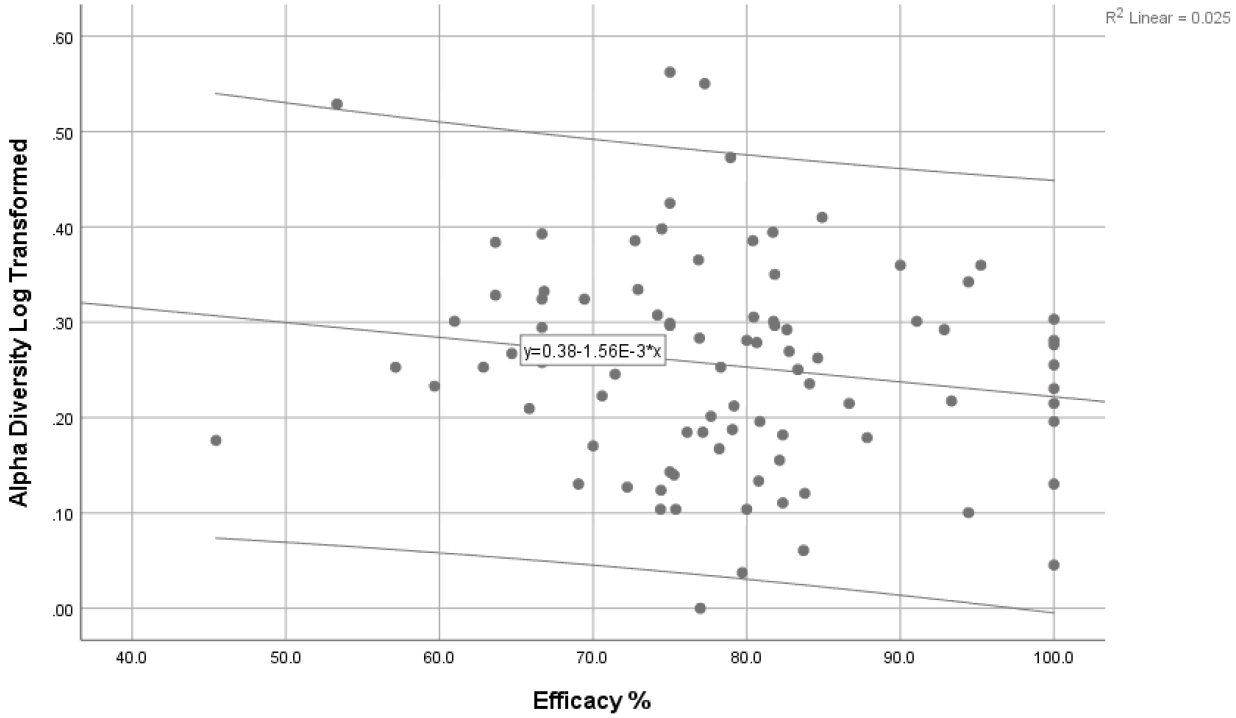




Appendix V. Post-transformation scatterplot for alpha diversity, log-transformed and clinical efficacy



Appendix W. Scatterplot of alpha diversity, log-transformed demonstrating 95% confidence intervals



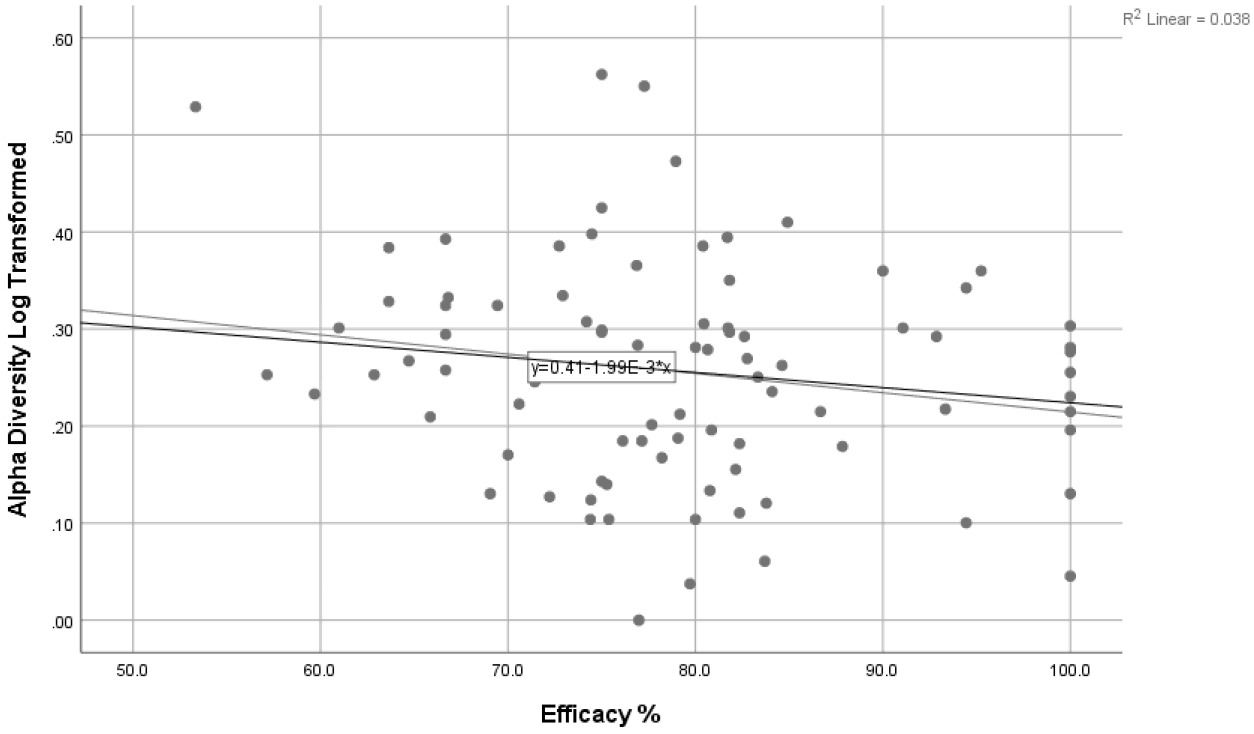
Appendix X. Casewise diagnostics to evaluate for potential outliers

Casewise Diagnostics[a]

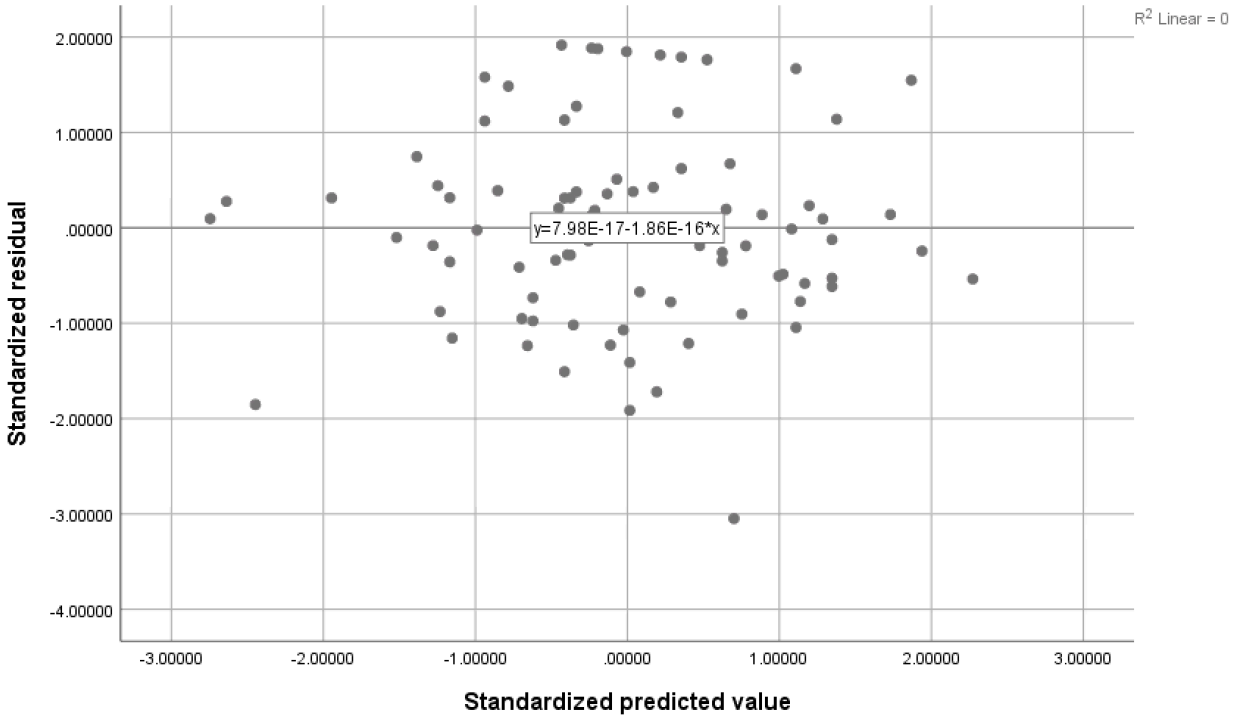
Case Number	Std. Residual	Efficacy %	Predicted Value	Residual
72	-3.049	45.5	80.219	-34.7685

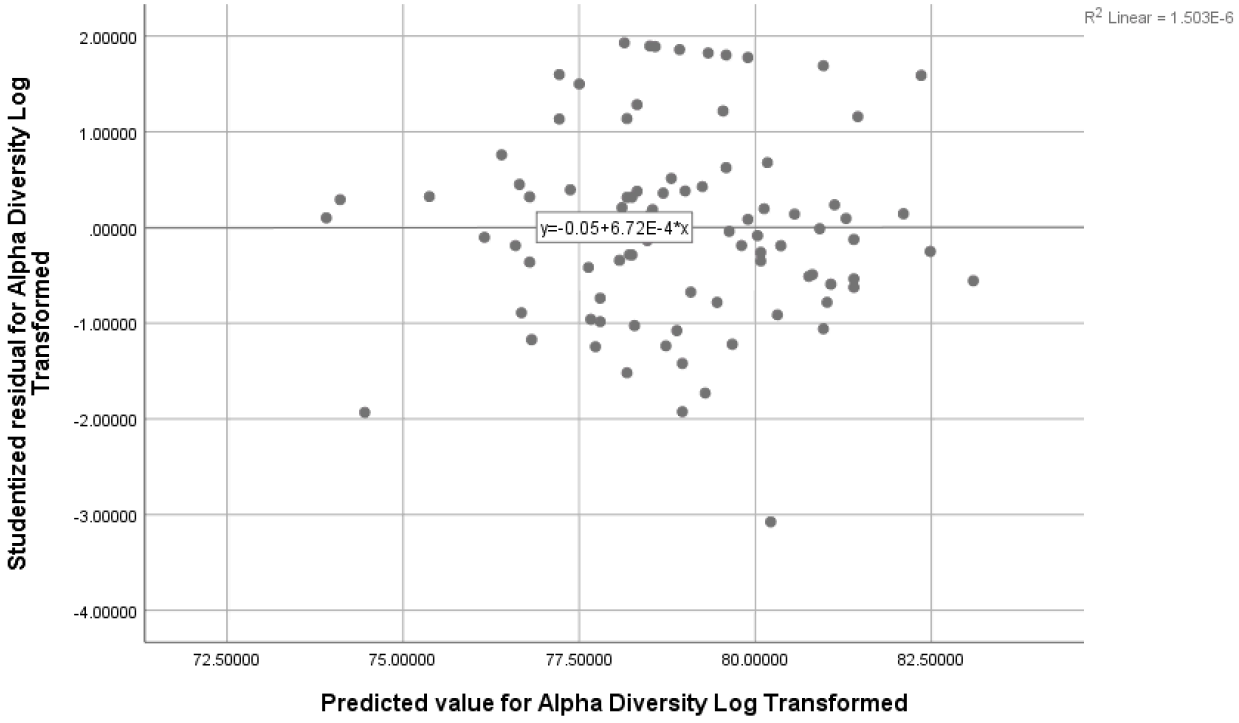
a Dependent Variable: Efficacy %

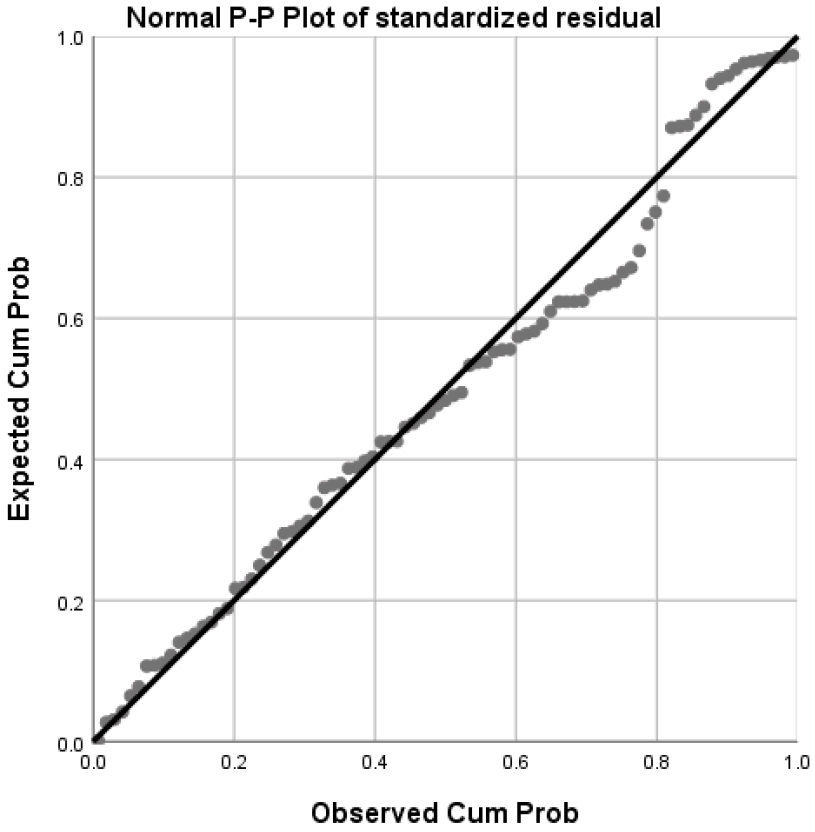
Appendix Y. Scatterplot for alpha diversity log-transformed post removal of suspected outlier (n=86)



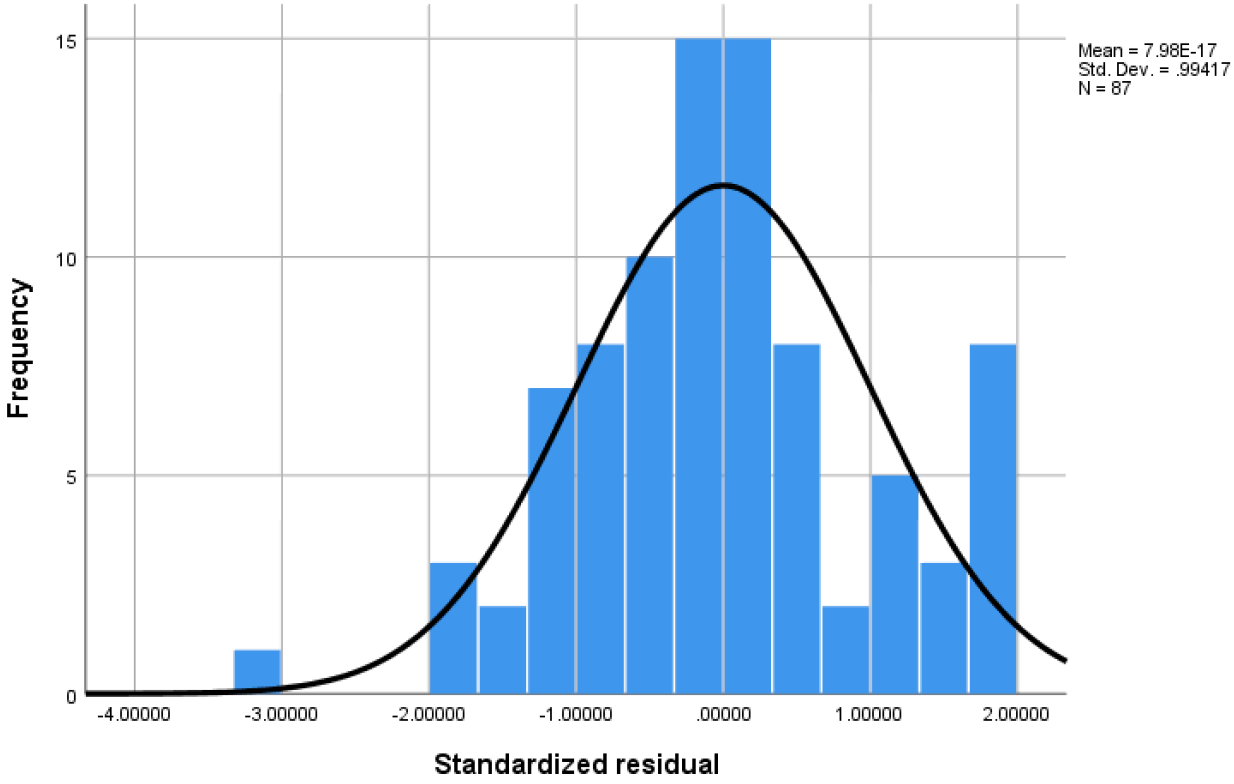
Appendix Z. Scatterplots and P-P plots of residuals for alpha diversity, log-transformed






Normal P-P Plot of standardized residual
Expected Cum Prob
Observed Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0

Appendix AA. Histogram of standardized residual for alpha diversity, log-transformed



Appendix BB. Results from model 1 for the Aim 2 simple linear regression

Descriptive Statistics

Variable	N	Mean	Std. Dev.
Efficacy %	87	78.9	11.5
Alpha Diversity log-transformed	87	0.255	0.1

Correlations

		Efficacy %	Alpha Diversity log-transformed
Pearson Correlation	Efficacy %	1.0	-0.159
	Alpha Diversity log-transformed	-0.159	1.0
Sig. (1-tailed)	Efficacy %		0.07
	Alpha Diversity log-transformed	0.07	
N	Efficacy %	87	87
	Alpha Diversity log-transformed	87	87

Variables Entered / Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Alpha Diversity log-transformed[b]		Enter

a Dependent Variable: Efficacy %
b All requested variables entered.

Model Summary[b]

Model	R	R Square	Adj R Square	Std. Error of the Estimate
1	0.159[a]	0.025	0.014	11.405

a Predictors: (Constant), Alpha Diversity log-transformed
b Dependent Variable: Efficacy %

Model Change Statistics

Model	R Square Change	*F* Change	df1	df2	Sig. *F* Change	Durbin-Watson
1	0.025	2.217	1	87	0.140	0.020

ANOVA[a]

Model		Sum of Squares	df	Mean Square	*F*	Sig.
1	Regression	288.3	1	288.3	2.217	0.14[b]
	Residual	11,055.9	85	130.1		
	Total	11,344.2	86			

a Dependent Variable: Efficacy %
b Predictors: (Constant), Alpha Diversity log-transformed

Coefficients[a]

Model		Unstandardized Coefficients		Std Coefficients	*t*	Sig.	95% C.I. for Exp(B)	
		B	S.E.	Beta			Lower	Upper
1	(Constant)	83.095	3.050		27.245	0.000	77.031	89.159
	Alpha Diversity log-transformed	-16.337	10.973	-0.159	-1.489	0.140	-38.154	5.481

a Dependent Variable: Efficacy %

Casewise Diagnostics[a]

Case Number	Std. Residual	Efficacy %	Predicted Value	Residual
72	-3.061	45.5	80.301	-34.851

a Dependent Variable: Efficacy %

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Dev.	N
Predicted Value	72.128	82.202	78.935	1.936	87
Residual	-34.851	21.638	0.000	11.321	87
Std. Predicted Value	-3.516	1.687	0.000	1.000	87
Std. Residual	-3.06	1.900	0.000	0.994	87

a Dependent variable: Efficacy %