NEW ARTICLE

This article was posted on the Archives Web site as a New Article. New Articles have been peer reviewed, copyedited, and reviewed by the authors. Additional changes or corrections may appear in these articles when they appear in a future print issue of the Archives. New Articles are citable by using the Digital Object Identifier (DOI), a unique number given to every article. The DOI will typically appear at the end of the abstract.

The DOI for this manuscript is doi: 10.5858/arpa.2021-0636-OA

The print version of this manuscript will replace the New Article version at the above DOI once it is available.

© 2022 College of American Pathologists
Correlation of Clinical Severity With Stool Microbiome Changes in *Clostridioides difficile* Infection

Daniel Castañeda-Mogollón, MSc; Cody P. Doolan, MSc; Noah B. Toppings, BSc; Ranmalee Amarasekara, PhD; Thuy-Ann Tran, BSc; Dylan R. Pillai, PhD, MD

• **Context.**—*Clostridioides difficile* infection (CDI) is the world-leading cause of infectious nosocomial diarrhea and pseudomembranous colitis. Antibiotics are the first line of treatment against CDI despite the high likelihood of antibiotic failure and/or recurrence. More data are needed to correlate clinical variables with 16S rRNA microbiome profiles in CDI-infected patients.

**Objective.**—To determine the relationship(s) between a patient’s clinical factors and the stool bacteriome of CDI-positive patients and CDI-negative patients with diarrheal symptoms.

**Design.**—This study used stool samples and clinical data from 358 patients with nosocomial diarrhea, who were divided by their CDI diagnosis (CDI-negative: n = 180; CDI-positive; n = 178). The stool bacteriome was profiled by amplicon deep sequencing of the 16S rRNA gene, followed by correlating clinical data.

**Results.**—The stool bacteriome was significantly different by severity assessment regardless of CDI status. Phyla and species varied significantly by CDI diagnosis. Severity, defined as a serum white blood cell count greater than 15 cells/μL and/or a creatinine level greater than 1.5 mg/dL, correlated significantly with dysbiosis of the stool bacteriome profile of CDI-positive patients compared to CDI-negative patients. Serum white blood cell count was significantly higher in patients with bacterial dysbiosis, and high levels of creatinine were associated with low bacteriome diversity.

**Conclusions.**—Clinical severity of CDI influences the stool microbiome of infected patients. To date, this study has the largest data set comparing 16S rRNA microbiome profiles and clinical variables between CDI-infected and noninfected individuals.

Arch Pathol Lab Med. doi: 10.5858/arpa.2021-0636-OA

---

*Clostridioides difficile* (formerly known as *Clostridium difficile*) infection (CDI) is the world-leading cause of nosocomial antibiotic-associated diarrhea and pseudomembranous colitis.2 *Clostridioides difficile* can cause a wide spectrum of clinical conditions, which range from mild diarrhea to life-threatening complications.2,3 Risk factors include age (most individuals with CDI are older than 65 years), hospitalization in an intensive care unit, a clinical background of inflammatory bowel disease, the use of proton pump inhibitors,4,5 and exposure to antibiotics.4,6,7 In 2017, the United States reported an estimated 462 000 CDI cases and 20 500 in-hospital deaths,8 with more than USD $4.0 billion in excess medical costs.3 In 2021, Canada reported an estimated 37 900 CDI cases with a cost of USD $281 million.9 In addition, an annual incidence decrease has been noted in Canadian populations between 2016 and 2017,3 despite the predominance of the more lethal hypervirulent BI/027/NAP1 strain.10

Current treatment guidelines from the American Journal of Gastroenterology, the Canadian Paediatric Society, and the European Society of Clinical Microbiology and Infectious Diseases suggest vancomycin and metronidazole as the first-line therapy.11–13 Alternatively, the Infectious Diseases Society of America (IDSA) now recommends fidaxomicin or vancomycin over metronidazole for an initial CDI episode for both severe and nonsevere disease.14 Nevertheless, these approaches rely on the susceptibility of *C difficile* to these antibiotics despite the emergence of antibiotic-resistant strains;15,16 Some antibiotic-free alternatives focus on restoring the stool microbiome as an effective way to treat the disease, such as a fecal microbiome transplant (FMT).17,18 Previous studies have highlighted the effectiveness of this method over the use of vancomycin and other antibiotics.17,18

The emergence of *C difficile* strains that are resistant to second-generation fluoroquinolones, clindamycin, and tetracycline15,19 calls for treatment alternatives for CDI that restore the stool microbiome and prevent CDI recurrence.

---

Accepted for publication May 27, 2022.

Supplemental digital content is available for this article. See text for hyperlink.

From the Departments of Pathology & Laboratory Medicine (Castañeda-Mogollón, Doolan, Toppings, Amarasekara, Tran, Pillai) and Microbiology, Immunology, and Infectious Diseases (Castañeda-Mogollón, Doolan, Toppings, Pillai), Cumming School of Medicine, and Calvin, Phoebe & Joan Snyder Institute for Chronic Diseases (Castañeda-Mogollón, Doolan, Toppings, Pillai), University of Calgary, Calgary, Alberta, Canada; and Alberta Precision Laboratories, Diagnostic & Scientific Centre, Calgary, Alberta, Canada (Pillai).

The authors have no relevant financial interest in the products or companies described in this article.

Corresponding author: Dylan R. Pillai, PhD, MD, Departments of Pathology & Laboratory Medicine, Microbiology & Infectious Diseases, & Community Health Sciences, University of Calgary, Diagnostic & Scientific Centre, Room 1W-416, 9-3535 Research Rd NW, Calgary, AB T2L 2K8, Canada (email: drpillai@ucalgary.ca).
Several studies have profiled the stool microbiome to determine the microorganisms that play a key role during CDI, including colonization.20–26 Other studies have highlighted the importance of clinical factors on the outcome in CDI patients.27–30 While these studies provide insights derived from either clinical variables or microbiome profiles, these 2 factors have not been analyzed together. To get a more complete picture of CDI, the decision was made to analyze clinical variables and microbiome profiles concurrently. Here we hypothesize that severity episodes in terms of a high white blood cell (WBC) count or high levels of creatinine will display different stool bacteriome profiles by CDI diagnosis.

This study characterized the stool bacteriome profile of 358 diarrheic hospitalized patients through amplicon deep sequencing of the 16S rRNA gene. First, a method validation between these patients was performed by comparing its profile to that of nonsymptomatic individuals. Secondly, we compared the relative abundance of each phylum in each patient by CDI diagnosis. Bacteriome diversity and overall profiles were used for severity assessment and correlation to various clinical factors, including WBC count, creatinine, albumin, and human C-reactive protein (CRP).

**METHODOLOGY**

**Ethics Approval**

Ethical approval was obtained through University of Calgary’s Institutional Research Information Services Solution, Calgary, Alberta, Canada (REB-18-0397). Information was protected as per Alberta Health Services protocol and data were de-identified before analysis.

**Sample and Population Characteristics**

Samples in this study were collected at Alberta Precision Laboratories in Calgary, Alberta, Canada. Briefly, 358 hospitalized patients presenting with diarrhea after hospital admission were screened for *C difficile* by using a 2-step testing algorithm as per Alberta Precision Laboratories guidelines (Figure 1). A stool sample was collected from the hospitalized patients and sent for analysis. An aliquot of these samples was used for this study before CDI treatment. Patients received one or several antibiotics including carbapenems, penicillins, quinolones, cephalosporins, tetracyclines, macrolide derivatives, glycopeptides, lincomycin derivatives, urinary tract anti-infectives, sulfonamides, aminoglycosides, oxazolidinones, or miscellaneous antibiotics, and were enrolled in the study between January 2019 and May 2020. Baseline characteristics collected from the study population were sex, age, presence of diarrhea, and antibiotic(s) administration. Stool samples from 9 non-diarrheic individuals were used as healthy controls for bacteriome profiling. Blood tests were run on hospitalized patients to assess the levels of WBC, creatinine, albumin, and CRP. Patients presenting with leukocytosis (WBC >15 cells/μL) or creatinine levels above 1.5 mg/dL were considered to be undergoing an episode of severe CDI, as described by the IDSA.33 Patients with serum albumin levels below 25 g/L were considered to be presenting with hypoalbuminemia as discussed elsewhere.32 Levels of CRP above 30 mg/L were considered high, as previously described.33 These levels were also used for severity assessment amongst the CDI-negative patients. No blood tests were performed on the healthy control cohort.

**Clostridioides difficile Testing**

Clostridioides difficile diagnosis was performed either by chemiluminescent immunoassay to detect glutamate dehydrogenase (GDH) specific for *C difficile* or by the Quik Chek Complete rapid membrane enzyme immunoassay for GDH antigen and toxin detection. If the original result from either test was negative, then patients were negative for *C difficile*. If either test result was positive, then the Cepheid GeneXpert (Sunnyvale, California) Dx System (a real-time polymerase chain reaction [PCR] test for rapid detection of toxin B, binary toxin, and the TcdC deletion sequences) was used to confirm the *C difficile* diagnosis; cycle threshold (Ct) values of 40 or greater were considered negative. If the results from the GeneXpert Dx System were negative, then the patient was considered negative for *C difficile* colonization.
tion. Assessment of the ribotype 027 strain was also included in the bacteriome diversity analysis. CDI-negative patients with diarrhea were not tested for other etiologies. Noninfectious nosocomial diarrhea etiologies were not investigated.

DNA Isolation, and 16S rRNA v3-v4 Amplification

Patient stool samples were collected and stored at −80°C until aliquoting. Next, stools samples were aliquoted and stored at −80°C until further use. Aliquots were then thawed, and then DNA was isolated by using the QIAamp DNA minikit (catalogue No. 51306) following the manufacturer’s protocol. The variable regions 3 and 4 of the 16S rRNA gene were amplified by using the Illumina 16S recommended primers with its adapters34 (forward: 5′-TCGTCGACGCTACGGGNGGCWGCAG-3′; reverse: 5′-GTCTCGTGGGCTCGAGATGTGTATAAGAGACAG-GACTACHVGGGTATCCTAATCC-3′). The first PCR consisted of 1 μmol of each primer, 5 μL of template DNA, and 11 μL of PlatinumTM SuperFi TM PCR Master Mix (catalogue No. 12358010, ThermoFisher Scientific). The thermocycling protocol used was as follows: 95°C for 3 minutes, followed by 30 cycles of 95°C for 30 seconds, 59°C for 30 seconds, and 72°C for 30 seconds, with a final 75°C for 5 minutes. The presence of a PCR product was verified by electrophoresis (2% agarose gel). Fragments were purified with AMPure XP beads (Beckman Coulter). This cleanup was repeated after the second PCR.

Library Preparation and Sequencing

16S rRNA gene libraries were prepared and sequenced as previously described.34 Briefly, 16S rRNA purified amplicons were dually barcoded by using the Nextera XT kit.34 Library size validation was done by capillary electrophoresis using the Agilent 5200 Fragment Analyzer System. DNA reads were sequenced with the Illumina MiSeq instrument (San Diego, California) with 2x300 in paired-end mode. A 5% spike of the PhiX enterophere was used as an internal sequencing control.

16S v3-v4 rRNA Sequence Processing and Primary Data Analysis

Read preprocessing was performed by using the QIIME 2.0 pipeline.35 Briefly, Deblur36 was used for denoising, removing low-quality reads, and truncating reads to 240 bp. VSEARCH was used for the removal of chimeras and borderline chimeras from PCR amplification errors in de novo mode.37 The machine-learning classify-sklearn38 plugin was used to assign taxonomies to the denoised sequences by using the pretrained Greengenes database39 with 99% operational taxonomic unit (OTU) similarity. This classifier has trimmed sequences including 250 bases from the 16S rRNA v4 region, to determine the taxonomy to the preprocessed sequences. Multiple sequence alignments were performed in de novo mode by using the MAFFT40 plugin. Alpha and beta diversity analysis were carried out by using the qiime diversity plugin. Shannon and Simpson indices were calculated for the alpha diversity analysis, and principal coordinate analysis (PCoA) plots were computed by using the Bray-Curtis dissimilarity metric for the beta diversity analysis of dissimilarity. The adレx21 plugin was used for analysis of differential abundance amongst OTUs between population cohorts. The sequences from the significantly abundant OTUs by cohorts in the volcano plot were individually assessed for species identification. For this, each sequence was aligned against the nucleotide database, using MegaBLAST.42 Each species was annotated only if the percentage of similarity between the query and the output was above 99% with an e-value below 1 × 10−100. Individuals with a Shannon index below 3.5 were considered to have dysbiosis (imbalance in the stool bacteriome).

Statistical Analysis and Visualization

Statistical analysis of the alpha diversity indices was conducted by using the 2-sided nonparametric Kruskal-Wallis or Mann-Whitney test for nonpaired comparisons. The beta diversity analyses were performed by plotting a PCoA using the Bray-Curtis dissimilarity metric, followed by a permutational multivariate analysis of variance (PERMANOVA) with 999 permutations. A pairwise comparison and P value adjustment using the Benjamini-Hochberg correction was done across the PCoAs comprising more than 2 clusters. Volcano plots were generated by using the log−fold change of specific OTUs across 2 cohorts, followed by its correspondent −log10(q-value) from a Mann-Whitney test, and adjusted with the Benjamini-Hochberg correction. Simple linear regressions and Pearson coefficients were also computed for continuous variables. A receiver operating characteristic (ROC) curve was generated to assess the analytical sensitivity and specificity of QIIME 2.0 as a diagnostic tool for C difficile colonization. Further analysis and visualization were performed with GraphPad Prism version 9.2.0 for Mac (San Diego, California) along with the ggplot2 v 3.5.5,44 qiime2R v 0.9.9.6, tidyverse v 1.3.1,44 and aledx v 1.16.041 packages in RStudio v 1.2.1355.45 P < .05 was considered significant for all the analyses.

RESULTS

Population and Clinical Characteristics

To the best of our knowledge at the time of writing, this is one of the first studies to investigate both the stool microbiome (via amplicon deep sequencing), and clinical variables of symptomatic hospitalized individuals. The number of enrolled patients is higher than that of the current studies correlating microbiome profiles with clinical factors. In this study, a total of 358 patients were enrolled during the study period. The demographics, clinical characteristics, and bacteriome features are detailed in the Table. The median age of the hospitalized patients was 64 years (interquartile range, 54–74 years) with no significant difference (2-tailed t test, \( P = .99 \)). The sex distributions of each cohort were not significantly different (2-tailed t test, \( P = .82 \)). No significant difference was found between the age of the CDI-positive and CDI-negative patients (Dunn pairwise comparison, adjusted \( P = .99 \)); however, the age of the control cohort and each hospitalized cohort was significantly different (Dunn pairwise comparison, adjusted \( P < .001 \)). Of 342 patients with WBC counts, 151 (44.15%) presented with leukocytosis (WBC > 15000 cells/μL). Elevated serum creatinine levels (>1.5 mg/dL) were observed in 142 of 340 patients (41.76%). Sixty-three of 261 patients tested (24.13%) presented with hypoalbuminemia (<25 g/L). High levels of CRP (>50 mg/L) were observed in 141 of 144 patients tested (98.60%). A severity assessment of hospitalized patients was performed on the basis of presence of leukocytosis and/or high serum creatinine levels. Severe CDI was identified in 121 of the 177 CDI-positive patients (68.36%) on the basis of leukocytosis and/or high serum
creatinine levels. Severity, regardless of \textit{C difficile} diagnosis, was identified in 216 of 342 patients (63.15%). No correlation was found between hypoalbuminemia and severity by leukocytosis and/or high creatinine levels (OR = 1.23; Fisher exact test, \(P = .54\)). Similarly, no correlation was observed between the presence of \textit{C difficile} and hypoalbuminemia (OR = .97; Fisher exact test, \(P > .99\)). CDI-negative patients were treated with antibiotics on the basis of their diagnosis. CDI-positive patients were treated with vancomycin, metronidazole, and/or piperacillin/tazobactam, in addition to other antibiotics used to treat any other comorbidities. No significant difference was found in confounding factors among the hospitalized individuals.

### Clostridioides difficile Infection Diagnosis

Of the 358 patients with diarrhea, 178 (49.72%) tested positive for \textit{C difficile} and 180 (50.27%) tested negative (Figure 1). Of the 358 patients with diarrhea, 53 (14.80%) were initially tested for \textit{C difficile} by using the Quik Chek Complete assay (which identifies the GDH antigen and/or \textit{C difficile} toxin). Of these 53 patients, 46 (86.79%) tested positive by GDH antigen and/or toxins A and B (Supplemental Table 1, see the Supplemental Digital Content containing 1 table and 9 figures). Among the 46 patients with positive findings by the Quik Chek Complete assay, 17 (36.96%) tested positive for GDH and toxins, and 29 (63.04%) tested positive for toxins but positive for GDH. Six patients among the 46 testing positive by Quik Chek Complete (13.04%) were PCR negative for the TcdB gene GeneXpert assay, and therefore assessed as negative for CDI. A total of 138 patients tested positive for the GDH chemiluminescent immunoassay and TcdB PCR (GeneXpert). Eighteen patients tested negative for TcdB PCR assay but positive for GDH (13.04%) and were assessed as negative for CDI. Nine patients among the 178 testing positive by TcdB PCR (5.08%) were also positive for the single nucleotide deletion at the TcdB gene, suggesting the presence of the more toxigenic 027/NAP1/BI \textit{C difficile} strain. Ct values by the TcdB gene and the TcdC single base deletion had similar values except for 1 patient (Supplemental Figure 1). Of the 358 individuals, 269 outcomes (75.14%) were recorded, while the rest (24.86%) were lost to follow-up. Of these 269 recorded outcomes, a total of 31 patients (11.52%) died within the facility. A total of 18 of 31 (58.06%) were deemed CDI positive, while the rest (41.94%) were considered CDI negative. No mortality significant differences were recorded by CDI group (Fisher exact test, \(P = .35\)).

### Bacteriome Diversity Differs by Clostridioides difficile Diagnosis and Presence of Diarrhea

A total of 365 samples were successfully sequenced. From this, 15 256 994 DNA reads were obtained. Of these, 8 243 518 reads belong to the CDI-positive group, 6 571 946 reads belong to the CDI-negative group, and 441 530 belong to the healthy control group. Two CDI-positive stool samples were of insufficient volume to begin extraction. Cohort-averaged OTU richness was assessed

#### Table: Demographics, Clinical Characteristics, and Bacteriome Features of the Study Subjects in Each Cohort

<table>
<thead>
<tr>
<th>Variables</th>
<th>\textit{C difficile} Positive (n = 178)</th>
<th>\textit{C difficile} Negative (n = 180)</th>
<th>Control Subjects (n = 9)</th>
<th>(P^{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean ± SD, y (n = 343)</td>
<td>63.58 ± 15.63 (n = 177)</td>
<td>63.96 ± 16.81 (n = 166)</td>
<td>29.8 ± 9.85 (n = 9)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Sex (n = 343)</td>
<td>177</td>
<td>166</td>
<td>9</td>
<td>.98</td>
</tr>
<tr>
<td>Male (n = 164)</td>
<td>84</td>
<td>80</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Female (n = 179)</td>
<td>93</td>
<td>86</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Blood tests</th>
<th>\textit{C difficile} Positive (n = 178)</th>
<th>\textit{C difficile} Negative (n = 180)</th>
<th>Control Subjects (n = 9)</th>
<th>(P^{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td>WBC, mean ± SD, cells/µL (n = 342)</td>
<td>16.48 ± 10.43 (n = 176)</td>
<td>14.62 ± 8.33 (n = 166)</td>
<td>N/A</td>
<td>.12</td>
</tr>
<tr>
<td>Creatinine, mean ± SD, mg/dL (n = 340)</td>
<td>223.17 ± 262.77 (n = 175)</td>
<td>195.63 ± 218.94 (n = 165)</td>
<td>N/A</td>
<td>.98</td>
</tr>
<tr>
<td>Albumin, mean ± SD, g/l (n = 261)</td>
<td>29.87 ± 8.42 (n = 134)</td>
<td>29.33 ± 6.72 (n = 127)</td>
<td>N/A</td>
<td>.67</td>
</tr>
<tr>
<td>CRP, mean ± SD, mg/L (n = 144)</td>
<td>121.24 ± 100.04 (n = 81)</td>
<td>77.67 ± 76.08 (n = 63)</td>
<td>N/A</td>
<td>.004</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clinical severity assessment (n = 342)</th>
<th>\textit{C difficile} Positive (n = 178)</th>
<th>\textit{C difficile} Negative (n = 180)</th>
<th>Control Subjects (n = 9)</th>
<th>(P^{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe (n = 216)</td>
<td>177</td>
<td>165</td>
<td>N/A</td>
<td>.04</td>
</tr>
<tr>
<td>Nonsevere (n = 126)</td>
<td>56</td>
<td>70</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WBC count (n = 356)</th>
<th>\textit{C difficile} Positive (n = 178)</th>
<th>\textit{C difficile} Negative (n = 180)</th>
<th>Control Subjects (n = 9)</th>
<th>(P^{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td>WBC &gt; 15 cells/µL (n = 165)</td>
<td>84</td>
<td>81</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>WBC ≤ 15 cells/µL (n = 191)</td>
<td>92</td>
<td>99</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Creatinine level (n = 340)</th>
<th>\textit{C difficile} Positive (n = 178)</th>
<th>\textit{C difficile} Negative (n = 180)</th>
<th>Control Subjects (n = 9)</th>
<th>(P^{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creatinine &gt; 1.5 mg/dL (n = 210)</td>
<td>105</td>
<td>105</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Creatinine ≤ 1.5 mg/dL (n = 130)</td>
<td>70</td>
<td>60</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bacteriome diversity</th>
<th>\textit{C difficile} Positive (n = 178)</th>
<th>\textit{C difficile} Negative (n = 180)</th>
<th>Control Subjects (n = 9)</th>
<th>(P^{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed OTUs, mean ± SD</td>
<td>134.1 ± 51.8</td>
<td>197.2 ± 98.9</td>
<td>302 ± 65.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Shannon index, mean ± SD</td>
<td>5.0 ± .9</td>
<td>5.5 ± 1.0</td>
<td>6.2 ± .4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Simpson index, mean ± SD</td>
<td>.93 ± .06</td>
<td>.94 ± .05</td>
<td>.98 ± .006</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>\textit{C difficile} reads, mean ± SD</td>
<td>77.6 ± 198.8</td>
<td>2.7 ± 22.7</td>
<td>.2 ± .7</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Abbreviations: CRP, C-reactive protein; N/A, not applicable; OTU, operational taxonomic unit; WBC, white blood cell.

\(a\) Fisher exact tests were performed for discrete variables. Two-tailed Student t test and Mann-Whitney test were performed for continuous variables.

\(b\) Bolded \(P\) values are deemed significant at a significance level below .05.

\(c\) Clinical severity refers to a WBC count > 15 cells/µL and/or creatinine level > 1.5 mg/dL.

Stool Microbiome \textit{C difficile} Infection—Castañeda-Mogollón et al
with rarefaction curves by looking at the mean of 10 permutations per point (Figure 2, a). Individual sample rarefaction curves were also generated (Supplemental Figure 2). The CDI-positive cohort displayed the smallest averaged number of observed OTUs, followed by the CDI-negative cohort, while the healthy cohort displayed the highest averaged number of observed OTUs (Figure 2, a). A total of 4945 OTUs were observed amongst the 3 cohorts. From this, 2132 (43.1%) were shared by the hospitalized cohorts regardless of the C difficile diagnosis (Figure 2, b). From the 4945 OTUs, only 28 (.56%) were unique in the healthy cohort, which can be explained by the lower sample size in this group. The alpha diversity Simpson index was significantly different between all pairwise permutations of the 3 cohorts (Kruskal-Wallis, P < .001, followed by Dunn multiple comparison tests [healthy versus CDI-positive, adjusted P < .001; healthy versus CDI-negative, adjusted P < .001; CDI-positive versus CDI-negative, P = .04]) (Figure 2, c). The alpha diversity Shannon index was also significantly different between all pairwise permutations of the 3 cohorts (Kruskal-Wallis, P < .001, followed by Dunn multiple comparison tests [healthy versus CDI-positive, adjusted P < .001; healthy versus CDI-negative, adjusted P < .001; CDI-positive versus CDI-negative, P < .001]) (Figure 2, d). The beta diversity analysis from the PCoA shows that clusters are significantly distant by cohort (Figure 2, e) (PERMANOVA, P = .001) and by its pairwise correction (P = .001 for all 3 comparisons).

**Phyla and Species Stratification by Clostridioides difficile Diagnosis**

Overall, the QIIME 2.0 pipeline was able to detect C difficile in most of the cases among the CDI-positive patients (67.04% presented at least 1 DNA read) and did not detect any in most of the CDI-negative patients (90% did not present any C difficile DNA reads) (Figure 3, a). In addition, 4 samples of the 24 (16.67%) that were initially positive by GDH of Quik Chek but negative by Cepheid GeneXpert

---

**Figure 2.** Clostridioides difficile–positive patients have lower diversity than the C difficile–negative patients and healthy controls. a, Rarefaction curve of the averaged observed diversity (OTUs) across cohorts by sampling read depth (mean of 10 permutations per point). b, Venn diagram displaying the total number of OTUs across the C difficile–positive (n = 176), C difficile–negative (n = 180), and healthy (n = 9) individuals. c, Violin plots of the alpha diversity Simpson index by C difficile diagnosis and analyzed by the nonparametric Kruskal-Wallis test with a pairwise Dunn test comparison. d, Violin plots of the alpha diversity Shannon index by C difficile status and analyzed by the nonparametric Kruskal-Wallis test with a pairwise Dunn test comparison. e, Bray-Curtis beta diversity PCoA plot by C difficile diagnosis and analyzed by a PERMANOVA test with 999 permutations. *Adjusted P < .05. **Adjusted P < .001. Abbreviations: OTU, operational taxonomic unit; PC, principal component; PCoA, principal coordinate analysis; PERMANOVA, permutational multivariate analysis of variance.
Figure 3. Phyla and species stratification by *Clostridioides difficile* diagnosis. a, *Clostridioides difficile* difficile diagnosis along with a normalized relative frequency phyla plot. b, Side-by-side dot plot with individual data points grouped by C difficile diagnosis and healthy states and analyzed by the nonparametric Kruskal-Wallis test with a pairwise Dunn test comparison. c, Normalized phyla relative proportion bar plot subclassified by *C difficile* diagnosis. Each phylum was evaluated by the nonparametric Mann-Whitney test. d, Volcano plot of microorganisms’ abundance by *C difficile* diagnosis. The dashed line represents the adjusted P cutoff for significance (1.30). Red dots represent microorganisms above the significance cutoff. e, Bar plot of the microorganisms whose abundance was significantly different with greater abundance across *C difficile*-positive (red) and *C difficile*-negative (green) patient groups. **P < .01. ***P < .001. Abbreviation: ns, not significant.
Bacteriome Profile Reveals a Clinical Association With CDI Patients

A significant association was found between the severity index by WBC count or creatinine levels and the presence of C difficile (OR = 1.59; Fisher exact test, P = .04). CRP was higher amongst CDI-positive patients (Mann-Whitney test, P = .004); however, no significant difference was observed by WBC count (Mann-Whitney test, P = .12), creatinine (Mann-Whitney test, P = .98), or albumin (Mann-Whitney test, P = .67) (Figure 4, a). No correlation was observed between either the Simpson or the Shannon index and WBC count (Supplemental Figure 7, a and b), albumin (Supplemental Figure 7, c and d), or CRP (Supplemental Figure 7, e and f). Nevertheless, the Simpson index showed a significant negative correlation with the creatinine values (Pearson r = -.16; P = .003) (Figure 4, b). Interestingly, no correlation was observed between the Shannon index and creatinine (Supplemental Figure 7, g). WBC count was higher amongst patients with dysbiosis regardless of their C difficile diagnosis (Mann-Whitney test, P < .001), but not by creatinine (Mann-Whitney test, P = .18), CRP (Mann-Whitney test, P = .55), or albumin (Mann-Whitney test, P = .41) (Figure 4, c). An alpha and beta diversity analysis was performed across patients by C difficile diagnosis and assessment of severity by high WBC count or high creatinine levels. A negative correlation was observed between the TcdB Ct value and the WBC count (Figure 4, d), suggesting higher levels of WBC with initial C difficile bacterial load. There were significant differences between the Shannon indices of severe–CDI–positive and nonsevere–CDI–negative patients (Dunn pairwise correction test, adjusted P = .002), as well as between the severe–CDI–positive and severe CDI–negative patients (Dunn pairwise correction test, adjusted P = .02) (Figure 5, a). Interestingly, Simpson indices were only significantly different between the severe–CDI–positive and nonsevere–CDI–negative patients (Dunn pairwise correction test, adjusted P = .04) (Figure 5, b). The beta diversity analysis revealed significantly distinct centroids by severity status (PERMANOVA, P = .04) (Supplemental Figure 8). A beta diversity–stratified analysis by severity and C difficile diagnosis was also performed, suggesting a significant difference overall (PERMANOVA, P = .001) (Figure 5, c). A pairwise comparison of this stratified plot was done to assess the statistical significance by severity and C difficile presence. The severe–CDI–positive patients were statistically different from the severe–CDI–negative patients (PERMANOVA, adjusted P = .001). The severe–CDI–negative patients were significantly different from the nonsevere–CDI–negative patients (PERMANOVA, adjusted P = .001). A borderline adjusted P value was observed between the severe–CDI–positive and the nonsevere–CDI–positive patients (PERMANOVA, adjusted P = .069). No significance was observed between the nonsevere–CDI–negative patients and the severe–CDI–negative patients (PERMANOVA, adjusted P = .64). An analysis of overlapping microbial prevalence at the species level was performed by severity status regardless of the C difficile diagnosis. Surprisingly, no microorganisms were significantly more abundant between cohorts (Supplemental Figure 9).

**DISCUSSION**

For individuals diagnosed with C difficile infection and treated with antibiotics, the likelihood of treatment failure is

**Arch Pathol Lab Med**

Stool Microbiome C difficile Infection—Castañeda-Mogollón et al 7
between 20% and 35%; amongst these, 40% to 60% will have a second recurrence.46 Despite this, American, Canadian, and European health guidelines recommend antibiotic usage as first line of treatment.11–13 FMT has been proven effective in the past for CDI treatment. However, there is more to be learned about the stool bacteriome profile changes and the relationship to clinical severity. In this study, we profiled the stool bacteriome of hospitalized individuals with and without CDI by sequencing the 16S rRNA gene of bacteria present in diarrheal stool samples. Correlations between the bacteriome diversity profile and clinical factors were identified. This approach has an advantage over existing studies in that it includes analyses of laboratory parameters that define severity of CDI such as WBC, creatinine, albumin, and CRP, which have been largely overlooked in the current *C difficile* microbiome literature.23,26,47–49

Major results include the reduction in bacteriome diversity and richness in CDI-positive patients compared to CDI-negative patients, and significantly different alpha diversity in CDI-positive patients stratified by severity status. The alpha diversity analysis reveals a different distribution of richness and evenness of the microbial community, whereas the beta diversity analysis shows a difference of microbial composition. The results showed that severity alone, regardless of CDI diagnosis, plays an important role in the stool bacteriome. Specifically, significantly different bacteriome profiles were observed by the beta diversity analysis and the alpha diversity Simpson index, but not by the Shannon index. The first metric gives more priority to the evenness of the bacteriome, while the latter gives more weight to overall diversity. This disagreement between alpha indices could suggest that severity alone relates more to the evenness of the stool bacteriome rather than the overall diversity. Interestingly, when looking at the beta diversity of the bacteriome of CDI-positive patients, borderline significance was observed ($P = .07$). This was not observed in CDI-negative patients ($P = .64$). In addition, patients with dysbiosis regardless of their CDI diagnosis were noted to have a higher WBC count than those without dysbiosis.

A negative correlation was observed between the Cepheid GeneXpert RT-PCR diagnostic test *TcdB* Ct values and the WBC count, suggesting that the *C difficile* bacterial load at the time of testing relates to the immune response. Nevertheless, the *TcdB* Ct values were not significantly different by severity status. Other studies have not found a correlation between severity and the burden of *C difficile*, which suggests that *TcdB* Ct value is not a good predictor of severity in CDI patients. CRP levels were higher in CDI-positive than CDI-negative patients, indicating a higher level of inflammation caused by *C difficile*, as widely reported in the past. The WBC count did not vary significantly between the CDI-positive and CDI-negative patients, suggesting this clinical factor alone is not enough for *C difficile* prediction, which support previous observations.52 Previous studies have correlated high levels of CRP with CDI-positive patients, which supports the findings here described.53 Interestingly, CRP has been used as a predictor to determine *C difficile* recurrence from the same
strain\textsuperscript{53} and mortality outcome from hospitalized patients.\textsuperscript{54} Creatinine levels were negatively correlated to Simpson indices, indicating a relationship between the bacteriome and the immune response from biomarkers related to kidney disease. A study performed by Liu et al\textsuperscript{55} suggests an association between the stool microbiome and patients with chronic kidney disease.

The analysis of the taxonomic composition of the stool bacteriome yielded similar results to those of previous studies using similar 16S rRNA amplicon deep sequencing approaches.\textsuperscript{23,51,56,57} Specifically, CDI-positive patients tend to have a less diverse and less rich microbiome than diarrheic CDI-negative and healthy individuals.\textsuperscript{23,51,56,57} Conversely, other studies have found no differences between the alpha and beta diversity amongst symptomatic individuals with and without CDI.\textsuperscript{26,58} Nevertheless, these studies have a lower sample size than the one presented here.

Not surprisingly, the analysis of the stool bacteriome of the hospitalized individuals at the phyla level showed that the majority were Firmicutes, Bacteroidetes, and Proteobacteria, which is supported by previous findings.\textsuperscript{26,59} Furthermore, the analysis of phyla revealed that the Firmicutes to Bacteroidetes ratio is significantly higher in patients with dysbiosis than in those with a richer and more diverse microbiome.\textsuperscript{60} In addition to these findings, previous studies have observed an increased Firmicutes to Bacteroidetes ratio in obese humans and mice versus their lean counterparts,\textsuperscript{61} whereas a decreased ratio was recorded in patients with inflammatory bowel disease.\textsuperscript{60}

At the species level, both Epulopiscium and Clostridioides species were significantly enriched in the CDI-positive cohort, suggesting they could potentially play an important role during the course of the infection, which has been previously described.\textsuperscript{62} Interestingly, Epulopiscium genus has been used in the past for the prediction of CDI cases across symbolic classification and logistic regression models.\textsuperscript{62} Conversely, some of the microorganisms identified in the CDI-negative cohort could be used to identify other infectious-nosocomial etiologies. Indeed, B fragilis, A putredinis, B vulgatus, and P excrementihominis have been associated with various disease states, including diarrhea and intra-abdominal abscesses, among others.\textsuperscript{63–68}

Potentially beneficial microbes were more abundant in CDI-negative patients, such as A muciniphila. This species has been reported to benefit the host by influencing glucose/lipid metabolism and promoting intestinal immunity, among others.\textsuperscript{69}
Gut microbiome profiles along with clinical data can provide some insight in determining the outcome of antibiotic-free therapies like FMT. For instance, donor selection based on microbiome composition and clinical history is the most significant factor influencing FMT outcome in CDI-positive patients.\textsuperscript{70,71} In addition, previous studies have predicted the need for repeated FMT, based on higher WBC count and low albumin levels.\textsuperscript{72} Other studies have found key transkingdom interactions that play a role in FMT success, suggesting microbiome profile screening of recipient and donor should be done to ensure a healthy microbiome restoration in CDI-positive patients.\textsuperscript{73} This suggests that a screening of microbiome profiles of donor and recipient, along with clinical variable data, could predict the outcome to a successful FMT intervention. Despite the high success rate of FMT in CDI-infected patients, this therapy is typically used as a last resort. Reasons for this include its low-efficient treatment option for gut dysbiosis, especially in chronic immune-mediated conditions, as well as the lack of standardization regarding dose and timing regimens.\textsuperscript{73}

Current studies are limited either by including only clinical data on CDI-positive individuals\textsuperscript{74–77} or by evaluating the 16S rRNA gut microbiome profiles with limited to absent clinical background.\textsuperscript{20,23,78–80} Studies including both are typically limited by their sample size.\textsuperscript{81,82} To our knowledge, this study comprises the largest data set to date comparing CDI-positive versus CDI-negative individuals while looking at 16S rRNA bacteriome data and clinical variables.

The empirical results reported herein should be considered in light of some limitations. Firstly, a small sample size was obtained for the control cohort (n = 9) to validate the bacteriome diversity from the hospitalized individuals. Secondly, a significant age difference was found between the aforementioned control cohort and the patients analyzed here, which could be a confounding factor when comparing the microbiome profiles across the hospitalized patients. The v3-v4 regions of the 16S rRNA gene did not always capture all the identified taxa at the species or strain level. Finally, no posttreatment follow-up data were available to consider the number of CDI relapses that could support the choice of FMT over antibiotics. Finally, the definition of clinical severity used in this study is somewhat arbitrary.

Severity measured by high levels of WBC and/or creatinine had a higher impact in CDI-positive patients than in patients suffering from nosocomial diarrhea caused by other etiologic agents. WBC count was significantly higher in patients with dysbiosis, and a negative correlation was observed between WBC and TcdB Ct values. CRP was significantly lower in the CDI-negative than the CDI-positive cohort, suggesting a higher level of inflammation in the latter. Creatinine was observed to correlate negatively with Simpson index, but not Shannon index, suggesting a higher correlation on the evenness of the microbiome than its overall diversity. No correlations between albumin and microbiome profiles were found. The combination of amplicon deep sequencing and clinical factors is a promising holistic tool for understanding CDI.

CONCLUSIONS

The overall diversity and richness of the stool microbiome was severely affected in CDI-positive patients compared to those with diarrhea caused by other etiologic agents. Severity, as determined by high levels of WBC or creatinine, had a higher impact in the stool microbiome amongst CDI-positive than CDI-negative patients. Higher levels of creatinine are associated with a less diverse stool microbiome. In addition, a higher Firmicutes to Bacteroidetes ratio was found amongst patients with dysbiosis and it was negatively correlated with the Simpson index. Similarly, a higher prevalence of overlapping species was identified in CDI-negative patients, including beneficial organisms like \textit{A. muciniphila}. More studies are needed to determine the association between Archaee and Eukarya in relation to stool diversity and clinical factors. These results provide a further understanding of how stool microbiome profiles are affected by clinical factors and they shed light on future studies into antibiotic-free therapies.

Data Availability

The 16S rRNA gene sequencing data generated by this study are available through the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under the accession number PRJNA770733.

We would like to thank the participants of the study. We would also like to thank the sequencing facility for their valuable input on the experimental design.

References


\textit{Stool Microbiome C difficile Infection}—Castañeda-Mogollón et al


79. Jenior ML, Leslie JL, Young VB, Schloss PD. *Clostridium difficile* alters the structure and metabolism of distinct cecal microorganisms during initial infection to promote sustained colonization. *mSphere*. 2018;3(3):e00261–18. doi:10.1128/mSphere.00261-18

