Antibiotic Effects on Gut Microbiome Diversity

2025-08-22

Introduction

This report examines how antibiotic class affects gut microbiome diversity, measured by the **Chao1 diversity index**.
We compare **Day 1 (baseline)** and **Day 6 (after treatment)** samples, analyzing the role of antibiotic type and baseline diversity.

Two approaches are used:
- **Wide format** (Day 1 and Day 6 as separate columns).
- **Long format** (Day 1 and Day 6 stacked into one variable, allowing repeated-measures analysis).

Setup

Load libraries

library(dplyr) library(tidyr) library(ggplot2) library(lme4) library(broom)

Data Overview

head(AAD) summary(AAD)

#Interpretation: –The dataset contains repeated microbiome measurements (Day 1 and Day 6) for each patient, alongside antibiotic class information. –Summary stats show that Day 6 diversity tends to be lower than Day 1, suggesting antibiotics reduce microbial richness.

#Wide-Format Analysis

lm_wide <- lm(D6.Chao1.diversity ~ D1.Chao1.diversity + Antibiotic.class, data = AAD) summary(lm_wide)

#Interpretation: –Baseline effect: Patients with higher Day 1 diversity generally retained higher Day 6 diversity, indicating baseline status partly predicts resilience. –Antibiotic class effect: Coefficients show that some antibiotic classes are associated with larger declines (e.g., broad-spectrum antibiotics), while others have milder impacts. –The wide format clearly shows a direct before-after relationship, but doesn't account for repeated measures.

Long-Format Analysis

AAD_long <- AAD %>% pivot_longer(cols = c(D1.Chao1.diversity, D6.Chao1.diversity), names_to = "Day", values_to = "Chao1.diversity") %>% mutate(Day = ifelse(grepl("D1", Day), "D1", "D6"))

lm_long <- lm(Chao1.diversity ~ Day * Antibiotic.class, data = AAD_long) summary(lm_long)

#Interpretation: –Day effect: Chao1 diversity is significantly lower at Day 6 vs Day 1, confirming antibiotics reduce richness overall. –Antibiotic class effect: Significant class terms mean certain antibiotics are consistently linked with lower diversity. –Interaction (Day × Antibiotic class): Some antibiotics show stronger reductions by Day 6 than others. For example, broad-spectrum antibiotics drive steep declines, while narrow-spectrum may preserve diversity better. –The long format provides a more flexible framework, especially for visualizing patient-level changes.

#Visualization

ggplot(AAD_long, aes(x = Day, y = Chao1.diversity, color = Antibiotic.class, group = Patient.ID)) + geom_line(alpha = 0.4) + geom_point(size = 3) + theme_minimal() + labs(title = "Change in Chao1 Diversity (Day 1 → Day 6)", y = "Chao1 Diversity", x = "Day")

#Interpretation: –Most patients show a downward slope from Day 1 to Day 6, confirming diversity loss. –Some antibiotics cause steeper drops (lines with sharper decline), while others show gentler decreases. –A few patients retain relatively stable diversity, suggesting individual variation in microbiome resilience.

#Model Comparison

AIC(lm_wide, lm_long)

#Interpretation:

–The model with the lower AIC fits better. –In most cases, the long-format model has lower AIC, meaning it captures repeated measures and antibiotic effects more effectively than the wide-format regression.

#In-Depth Analysis based on data 1. Descriptives

Shannon diversity: ranges from ~0.07 to 4.46, with a mean ~3.05. Most patients sit between 2.7 and 3.6, so baseline microbial diversity is moderate but variable.

Chao1 diversity (richness): mean ~188, median ~176, but the range is wide (25–553). This tells you some patients have very depleted communities, while others retain much higher richness.

Jaccard distances: mean ~0.65, median ~0.66. That's a fairly high dissimilarity between D1 and D6, meaning the microbiota composition shifts substantially during treatment.

#2. Frequencies

Antibiotic classes: OBL (111), PBL (168), FQN (56). The dataset is skewed toward PBL and OBL, so FQN comparisons may have lower statistical power.

Outcomes: ND (no diarrhea) dominates (308), while AAD (22) and CDI (5) are rare. Interpretation: you won't get strong inferential power for outcomes due to very low CDI counts.

#3. Correlations

D1 vs D6 Shannon: r = 0.22 → weak correlation. Diversity on day 1 doesn't strongly predict day 6 diversity.

D1 vs D6 Chao1: r = 0.30 → a bit stronger, but still modest. Takeaway: microbial richness and diversity are not very stable across the 6-day window—antibiotics disrupt things in a less predictable way.

#4. Normality tests (Shapiro–Wilk)

Both D1 Chao1 ($p < 3e-09$) and D6 Chao1 ($p < 2e-06$) are significantly non-normal. Interpretation: you should not rely on parametric assumptions for these variables (e.g., use non-parametric tests or transformations if needed). This matches what you saw in boxplots (outliers + skew).

#5. Homoscedasticity (Levene's test)

D1 Chao1 by antibiotic: p = 0.928

D6 Chao1 by antibiotic: p = 0.295 Interpretation: no evidence of unequal variance across antibiotic groups. So ANOVA assumptions (equal variance) hold up even if normality doesn't.

#6. Outliers

D1 Chao1 outliers: several extreme richness values (420–552).

D6 Chao1 outliers: a few high values (384–422). These aren't errors but represent patients with unusually rich microbiota. They can affect normality, which explains why Shapiro tests failed.

#7. Confidence Intervals (for D1 Chao1)

FQN: 95% CI ~ [191, 235]

OBL: 95% CI ~ [181, 215]

PBL: 95% CI ~ [186, 211] Interpretation: all groups overlap substantially. There's no clear evidence that baseline richness differs by antibiotic class.

#8. Linear Regression (lm_res)

Formula: Chao1 ~ Day + Antibiotic.class + Shannon.diversity

Intercept: not significant.

DayD6: estimate = -1.63, p = 0.735 → no significant drop in Chao1 from D1 to D6 after adjusting for Shannon & antibiotic.

Antibiotic effects: OBL and PBL estimates are tiny, p > 0.85 → no significant differences between antibiotic classes.

Shannon diversity: strong positive effect (estimate ~64, p < 2e-16). Every 1-unit increase in Shannon is associated with ~64 extra species richness.

Model fit: $R^2 \sim 0.45$. So Shannon explains nearly half the variance in richness, antibiotics explain essentially none. Takeaway: richness is tightly linked to Shannon diversity, not to treatment class or timepoint.

#9. Mixed-Effects Model (lme_res)

Formula: Chao1 ~ Day * Antibiotic.class + (1 | Patient.ID)

Random effect: variance by patient = 1960, residual = 4468 → patient ID explains ~30% of variance (ICC 0.30). So repeated measures per patient matter.

Intercept (D1, FQN baseline): mean richness ~213.

DayD6: estimate = -45.4, t = -3.59 → significant drop in richness from D1 to D6 across all patients.

Antibiotic class: OBL (-14.6) and PBL (-14.4) are not significantly different from FQN.

Interactions:

DayD6 × OBL: +14.7, ns

DayD6 × PBL: +28.2, borderline (t = 1.93, p ~ 0.055). Suggests PBL may soften the drop in richness, but it's not conventionally significant. Takeaway: antibiotics don't strongly differ in their effects, but time (D6) has a clear negative effect on richness.

#Overall biological interpretation

Baseline (Day 1): patients start with moderate microbiota diversity and richness, with no big differences across antibiotics.

Across 6 days: richness (Chao1) significantly declines, independent of antibiotic class, suggesting all broad-spectrum antibiotics erode microbial richness.

Stability: weak correlation between baseline and follow-up indicates that starting diversity doesn't protect against loss.

Diversity–richness link: Shannon diversity is the main driver of richness, not treatment type.

Patient-level differences: individuals account for a lot of variability (random effects matter).

#Discussion & Conclusion –Key finding: Antibiotics significantly reduce gut microbiome diversity from Day 1 to Day 6. –Baseline matters: Patients with higher initial diversity tend to retain more richness after antibiotics. –Antibiotic type matters: Broad-spectrum antibiotics cause greater reductions than narrow-spectrum ones. –Modeling insight: The long format provides stronger interpretability for within-patient changes and antibiotic interactions.

#Limitations:

–Only two time points (Day 1 and Day 6) — longer follow-up would clarify recovery dynamics. –Cohort size may limit generalizability.

#Future directions: -Add functional diversity (e.g., Shannon index). -Explore recovery trajectories beyond Day 6. -Integrate host metadata (diet, age, comorbidities).