# STATISTICAL ANALYSIS PLAN

Validation of eNAT medium for microbiome sequencing of fecal and oral samples

Matthew Kelly, Rebecca Roha Young

## Version History

| Version | Date | Reason for Update | Time Stamp |
|---|---|---|---|
| 1 | February 25, 2019 | Original Version | February 25, 2019 |
| 2 | | | |
| 3 | | | |

## BACKGROUND:

The human microbiota consists of the 10-100 trillion symbiotic microbial cells harbored by each person, primarily bacteria in the gut. Microbiome projects worldwide have been launched with the goal of understanding the impact of these symbionts on human health. The diversity among the microbiome of individuals is immense compared to genomic variation: individual humans are about 99.9% identical to one another in terms of their host genome, but can be 80-90% different from one another in terms of the gut microbiome. These findings suggest that employing the variation contained within the microbiome will be more fruitful in personalized medicine, the use of an individual patient's genetic data to inform healthcare decisions, than approaches that target the relatively constant host genome.

eNAT™ (Copan Italia, Brescia, Italy) is a non-nutritive transport medium intended for transport and preservation of clinical specimens to be analyzed by nucleic acid amplification techniques. eNAT cannot be used for culture-based assays because it contains a detergent and a protein denaturant able to completely inactivate microbial viability within 30 minutes for the majority of the tested microorganisms. eNAT is a maintenance medium having the ability to maintain the integrity of nucleic acids of cells or pathogens. eNAT is provided in tubes containing the medium, with or without swabs (FLOQSwabs™) for vaginal, endocervical, nasal, nasopharyngeal and buccal sample collection. eNAT can be also used for collection and storage of urine samples.

## STUDY OBJECTIVES:

This study is intended to compare the diversity and composition of the fecal and oral microbiota of samples stored in a standard medium (RNAlater) compared with eNAT medium. Furthermore, stability of the diversity and composition of the microbiota in fecal and oral samples stored in eNAT at room temperature will be evaluated.

## STATISTICAL HYPOTHESIS:

### Primary Hypotheses:

1.) There will be no difference in microbial alpha diversity (mean Shannon Diversity Index and Chao1 richness) of either fecal or oral samples prepared in A.) RNAlater and stored immediately at -80$^{O}$C and B.) eNAT and stored immediately at -80$^{O}$C.
2.) There will be no difference in microbial alpha diversity (mean Shannon Diversity Index and Chao1 richness) of either fecal or oral samples prepared in B.) eNAT and stored immediately at -80$^{O}$C and C.) eNAT and stored at room temperature for 30 ± 2 days.
3.) There will be no difference in overall composition of the microbiota of either fecal or oral samples prepared in A.) compared to B.), and also B.) compared to C.), as measured using Bray-Curtis dissimilarity matrices.

### Secondary Hypothesis:

1.) There will be no difference in relative abundances of genera between either fecal or oral samples prepared in A.) compared to B.), and also B.) compared to C.).

## STUDY POPULATION:

Fecal and oral samples were collected from a cohort of children undergoing hematopoietic stem cell transplantation (HSCT). Subjects were enrolled during the pre-HSCT evaluation or upon admission to the hospital for HSCT, and were followed until 100 days after HSCT. Fecal samples were collected on as many days as possible during the enrollment period. Oral swab samples from the right buccal mucosa were collected weekly from study participants.

## SAMPLE DETAILS:

50 fecal samples were collected from 13 participants. The minimum number of stool samples per patient was 1 and the maximum was 11. 27 saliva samples were collected from 6 participants. The minimum number of saliva samples per patient was 1 and the maximum was 8. Each sample was divided into three aliquots and the aliquots were prepared as follows:

A.) The sample was placed in RNAlater solution and immediately transferred to a -80$^{O}$C freezer;
B.) The sample was place in eNAT solution and immediately transferred to a -80$^{O}$C freezer;
C.) The sample was placed in eNAT solution and stored at room temperature for 30 ± 2 days prior to placement in a -80$^{O}$C freezer.

## BLINDING:

All analyses will be conducted on a dataset with a scrambled variable for sample preparation method. An individual not involved in the analysis will keep a data file with the unscrambled variable, and only when analysis decisions are finalized and statistical coding is complete will the statistician be unblinded.

## STATISTICAL ANALYSIS:

### Definition of Outcomes

The Shannon Diversity Index will be used to measure microbial diversity. The Chao1 Index will be used to measure species richness. The Shannon Diversity and Chao1 Indices will both be calculated using the estimate_richness function in the R phyloseq package.

Both the Shannon Diversity and Chao1 Indices are expected to be normally distributed. However, in the event that either is not normally distributed (Shapiro-Wilk statistic <0.97), a transformation will be performed to make the distribution approximately normal. The resulting statistics will subsequently be back-transformed.

### Effect Size Calculations

50 stool samples were collected; each sample was split into 3 aliquots (1 aliquot for each of the 3 preparation methods). Assuming a power of 0.8, a two-sided difference in means of paired samples, and a normal distribution of the Shannon Diversity Index (and Chao1 Index), we will be able to detect an effect size of 0.29. With 50 stool samples, and considering the 25 most abundant samples, an exponential distribution of the relative abundance of OTUs, and 80% power, a 16% change in the 3 most abundant OTUs could be detected (https://fedematt.shinyapps.io/shinyMB/).

27 saliva samples were collected for each preparation method. Using the same assumptions, we will be able to detect an effect size of 0.38 in the Shannon Diversity or Chao1 Indices.

### Statistical Methods

Microbiome Sequencing Data Pre-Processing

Sequences will be split, quality-trimmed, demultiplexed, and chimera-reduced with the use of QIIME tools. High-quality sequences will be clustered based on nucleotide sequence identity using DADA2, and given taxonomic assignments based on the alignment with the Greengenes database. Each read will be grouped into an operational taxonomic unit (OTU), a cluster of similar 16S rRNA sequences corresponding roughly to a species or a group of closely related species. Sparse OTUs with fewer than 10 counts across all samples or that do not appear in at least 10% of samples will be removed.

Primary Hypotheses:

The data follow a split-split-plot design, where there are three nested levels of experimental units. The first level is subject, the second level is the sample, and the third level is the aliquot. A mixed-model ANOVA model will be run (SAS, PROC MIXED), where the independent variable of interest is sample preparation method and the outcome is either Shannon Diversity Index or Chao1 Index (Primary Hypotheses 1 and 2). To account for the split-split-plot study design, a random effect of participant will be included, where sample is nested within participant. The random effect of sample will also be included, where aliquot is nested within sample. Each model will be run twice – once for fecal samples and once for oral swabs. The LS-Means and a 95% confidence interval will be presented.

Additionally, a Bland Altman plot will be used to describe the agreement between Shannon Diversity Index (and, separately, Chao1 Index) and pairwise sample preparation methods. A Bland-Altman plot (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4470095/) is an XY scatterplot, where the Y axis is the difference between the each paired Shannon Diversity Index and the X-axis is the average of the Shannon Diversity Index. If the sample preparation methods were to perfectly agree, the Bland-Altman plot would show a horizontal line where the difference between measurements is 0. If the distributions of either the Shannon Diversity or Chao1 Indices are not normal, the Bland Altman plot will be performed on transformed values.

For all parametric models, the residuals will be outputted and tested for normality. If the residuals from the models are not normally distributed (Shapiro-Wilk statistic <0.97), or the scatterplot of residuals vs. fit shows patterns or evidence of heteroscedasticity, the following options will be considered: 1.) outcome variables will be transformed and back-transformed statistics will be presented; or 2.) any outlier in the response variable will be winsorized to either the 95$^{th}$ percentile for high outliers or the 5$^{th}$ percentile for low outliers and a sensitivity analysis will be conducted showing both the unwinsorized result and the winsorized result; or 3.) if neither of these options leads to normally distributed residuals, the model will be repeated with a Wilcoxon rank-sum test. All hypothesis testing will be carried out at the 5% (2-sided) significance level unless otherwise specified.

Permutation analysis of variance (PERMANOVA) will be used to determine if there is a global difference in stool microbiota composition according to sample preparation method (Primary Hypothesis 3). The primary distance measure will be Bray-Curtis dissimilarity. 4,999 permutations will be used, and the random seed will be 3251.

Secondary (Exploratory) Hypothesis:

As an exploratory analysis, the R package metagenomeSeq be used to identify genera with differential relative abundance between preparation methods for stool samples only. The abundances will be aggregated at the genus level. Next, the phyloseq object will be converted to a metagenomeseq object and will then be normalized to the cumulative 75$^{th}$ percentile. The fitFeatureModel of the metagenomeSeq package will be used to fit a zero-inflated model for each specific genus separately. Models will first compare the difference between preparation methods A.) and B.), and will then evaluate for differences between preparation methods B.) and C.).

A second exploratory analysis will be completed for stool and oral samples except, instead of using metagenomeseq, a Wilcoxon signed-rank test will be used to evaluate for a difference in median relative abundance for the 20 most abundant genera. The difference in medians of these genera will be plotted and, for stool samples only, the results will be compared to those obtained using the zero-inflated models implemented using the R package metagenomeseq.

An FDR correction based on the Benjamini-Hochberg procedure will be employed in the analyses for both exploratory aims.

**Software Programs**

All analyses will occur either in R (with R Studio and the phyloseq and metagenomes packages) or SAS version 9.4.