STATISTICAL ANALYSIS PLAN

TUDY TITLE: MODULATION OF GUT MICROBIOTA WITH NBT-NM108 AS AN
ARLY TREATMENT FOR SUSPECTED OR CONFIRMED COVID-19
ATIENTS
HORT NAME: COVGUT20
ROTOCOL ID: NBTNM10810012020
CLINCIALTRIALS.GOV ID: NCT04540406

06.14.24

Study Analysis Plan

Sequencing Data Analysis

DNA, protein, and metabolites will be extracted from fecal samples. 16S rRNA gene V4 region sequencing and metagenomic shotgun sequencing with different platforms such as Ilumina and PacBio will be done using the extracted DNA. Metaproteome profiling will be achieved using mass spectrometry-based proteome analysis based on extracted protein. Metabolomics will be measured using fecal sample supernatant. Multi-omics data will be combined with clinical data for statistical analysis to explore the overall changes during COVID-19 course and changes induced by the treatment.

16S rRNA gene sequencing data will be analyzed based on the amplicon sequence variants (ASVs) [1] to determine the gut microbiota composition using QIIME 2 [2]. Shannon index and ASV richness will be used to assess alpha diversity. The principal coordinate analysis will be used to compare and visualize dissimilarity in gut microbiota structure between samples based on beta diversity (Weighted and Unweighted UniFrac distance). Statistical differences will be tested using permutational multivariate analysis of variance (PERMANOVA). PICRUSt 2 will be used to predict the functions of individual ASVs and the collective functions of the gut microbiota at the gene content and pathway levels [3]. Global functional profiles of the gut microbiota will be compared and visualized by principal component analysis.

Metagenomic sequencing data will be analyzed in a genome-centric way. The raw sequencing data will be processed with KneadData (https://huttenhower.sph.harvard.edu/kneaddata) for quality control. High-quality draft genomes will be de novo assembled from the high-quality sequencing data. The quality assessment, taxonomic assignment, and functional annotation of the genomes will be conducted by using CheckM [4], GTDB-Tk [5], and Prokka [6] respectively. The principal coordinate analysis will be used to compare and visualize dissimilarity in gut microbiota structure between samples based on beta diversity, and statistical differences will be tested using permutational multivariate analysis of variance (PERMANOVA).

The metaproteomes will be analyzed by using Unified Human Gastrointestinal Genome (UHGG) v2.0 as the reference for functional annotation and protein sequences database. The predicted protein sequences from metagenomic dataset will be aligned with the measured protein sequences from metaproteomic dataset for integrative analysis. Global level protein changes of the gut microbiota will be compared and visualized by principal component analysis. The metabolome will measure both polar metabolites and lipids. Global level of compound changes of will be compared and visualized by principal component analysis.

Repeat measures correlation will be used to assess the relationships between ASVs/high-quality draft genomes [7], followed by ASV/high-quality draft genome clustering based on co-abundance patterns. Multivariate methods such as MaAsLin2, that allow adjustment for

confounding variables and model the covariates as random effects, will be applied to interpret the relationship between ASVs / ASV co-abundance groups/functions/high-quality draft genomes/high-quality draft genome co-abundance groups and the clinical metadata/proteins/metabolites.

Clinical Data Analysis

To analyze the clinical data, Aalen's additive survival model will be used to test whether the impact of covarites on the cumulative hazard is time-dependent [8]. If the covariates do not vary with time, cox proportional hazards model will be subsequently used to adjust for age, gender, race/ethnicity, baseline COVID-19 severity, and whether the participant has taken treatments that may reduce COVID-19 severity (vaccination and monoclonal antibody treatment) as covariates. Survival analysis Kaplan-Meyer curves and log-rank test will be used to estimate and compare the risk of death and the risk of recovery, as a function of time, in participants with COVID-19-like symptoms [9, 10]. To compare the time to hospitalization, the time to recovery, the time to complete resolution of subjective symptoms, and the time to complete resolution of objective symptoms between treatment groups, we will use competing risks survival analysis (treating death or self-reported illness severity at Days 1, 14, 28 and 56 as a competing risk) [10, 11], use survival analysis without competing risk, and explore joint survival analysis and Bayesian survival model [12]. Specifically, cumulative incidence functions will be estimated and tested using Gray's test [13]. The Fine-and-Gray subdistribution hazard regression analysis [11] will be used to further adjust for age, gender, race/ethnicity, baseline COVID-19 severity, and whether the participant has taken treatments that may reduce COVID-19 severity (vaccination and monoclonal antibody treatment) as covariates. Proportion of participants who are "alive and not admitted to the hospital", proportions of participants who visit the emergency room at Days 1, 14, 28 and 56, have complete resolution of objective symptoms, have complete resolution of subjective symptoms and have complete resolution of subjective symptoms except cough and fatigue at Days 1, 14, 28 and 56 will be compared between groups using separate logistic regression analysis at each time point. Time for participants to have complete resolution of objective symptoms, complete resolution of subjective symptoms and complete resolution of subjective symptoms except cough and fatigue will be compared between groups using competing risks survival analysis described previously, treating death or self-reported illness severity at Days 1, 14, 28 and 56 as a competing risk [9, 11]. To compare the longitudinally (repeatedly) measured illness severity based on the Ordinal Scale for Clinical Improvement from the World Health Organization, body temperature (oral), oxygen saturation level, pulse rate, respiratory rate and fasting blood glucose between treatment groups, accounting for the effect of death, reported adverse event, or other appropirate competing risk, we will compare the treatment effect using the method of joint survival and longitudinal data analysis [14, 15], where each of the clinical

outcomes (e.g., illness severity, body temperature (oral), and oxygen saturation level, etc.) will be modeled using linear mixed models, and death, reported adverse event, or other appropirate competing risk will be modelled using Weibull proportional hazards model [16]. Age, gender, race/ethnicity, baseline COVID-19 severity, and whether the participant has taken treatments that may reduce COVID-19 severity (vaccination and monoclonal antibody treatment) will be controlled as covariates in these statistical analyses, where appropriate. Within-group changes of main outcomes between groups will be modeled as change from pre- to post-treatment usng ANOVA and modeled as the post-treatment value as the change itself with the pretreatment value entered as covariate using ANCOVA [17]. Lastly, alternative statistical methods will be explored to quantify the underlying differences between groups with respect to a time-to-event end point, robust methods include calculating ratio/difference of t-Year Survival Rates, ratio/difference of percentiles of survival functions, and ratio/differenceof restricted mean survival times or restricted mean time lost [18]. Statistical analyses described above will be conducted for both intent-to-treat dataset and per-protocal dataset. For each test of an outcome, we define the statistical significance by $p<0.05$. The false discovery rate [19] will be applied for multiple testing, where appropriate. Sensitivity analysis using multiple imputation methods [20] will be performed to handle the missing data. Other approaches, such as methods of selection models [21] or use the pattern-mixture models such as the control-based pattern imputation approach, or the tipping-point approach [22, 23] will also be considered.

Reference

1. Callahan, B.J., et al., *DADA2: High-resolution sample inference from Illumina amplicon data.* Nat Methods, 2016. **13**(7): p. 581-3.
2. Bolyen, E., et al., *QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science*. 2018, PeerJ Preprints.
3. Douglas, G.M., et al., *PICRUSt2: An improved and extensible approach for metagenome inference.* bioRxiv, 2019: p. 672295.
4. Parks, D.H., et al., *CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes.* Genome Res, 2015. **25**(7): p. 1043-55.
5. Chaumeil, P.A., et al., *GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database.* Bioinformatics, 2019.
6. Seemann, T., *Prokka: rapid prokaryotic genome annotation.* Bioinformatics, 2014. **30**(14): p. 2068-9.
7. Bakdash, J.Z. and L.R. Marusich, *Repeated Measures Correlation.* Front Psychol, 2017. 8: p. 456.
8. Martinussen, T., et al., *Estimation of direct effects for survival data by using the Aalen additive hazards model.* 2011. **73**(5): p. 773-788.
9. Putter, H., M. Fiocco, and R.B. Geskus, *Tutorial in biostatistics: competing risks and multi-state models.* Stat Med, 2007. **26**(11): p. 2389-430.
10. Meira-Machado, L., et al., *Multi-state models for the analysis of time-to-event data.* Stat Methods Med Res, 2009. **18**(2): p. 195-222.
11. Fine, J.P. and R.J. Gray, *A Proportional Hazards Model for the Subdistribution of a Competing Risk.* Journal of the American Statistical Association, 1999. **94**(446): p. 496-509.
12. Klein, J.P., et al., *Handbook of survival analysis*. Chapman & Hall/CRC handbooks of modern statistical methods. 2014, Boca Raton, FL: CRC Press.
13. Gray, R.J., *A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk.* The Annals of Statistics, 1988. **16**(3): p. 1141-1154.
14. Tsiatis, A.A. and M. Davidian, *JOINT MODELING OF LONGITUDINAL AND TIME-TO-EVENT DATA: AN OVERVIEW.* Statistica Sinica, 2004. **14**(3): p. 809-834.
15. Rizopoulos, D., *JM: An R package for the joint modelling of longitudinal and time-to-event data.* Journal of Statistical Software (Online), 2010. **35**(9): p. 1-33.
16. Rizopoulos, D., *JM: An R Package for the Joint Modelling of Longitudinal and Time-to-Event Data.* 2010, 2010. **35**(9): p. 33 %J Journal of Statistical Software.
17. French, A., et al., *Multivariate analysis of variance (MANOVA)*. 2008.
18. Uno, H., et al., *Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis.* J Clin Oncol, 2014. **32**(22): p. 2380-5.
19. Benjamini, Y. and D. Yekutieli, *The control of the false discovery rate in multiple testingunder dependency.* 2001. **29 %J The Annals of Statistics**(4): p. 1165-1188, 24.
20. Siddique, J., O. Harel, and C.M. Crespi, *Addressing Missing Data Mechanism Uncertainty using Multiple-Model Multiple Imputation: Application to a Longitudinal Clinical Trial.* Ann Appl Stat, 2012. **6**(4): p. 1814-1837.
21. Molenberghs, G. and G. Verbeke, *A review on linear mixed models for longitudinal data, possibly subject to dropout.* Statistical Modelling, 2001. **1**(4): p. 235-269.
22. Little, R.J.A., *Modeling the Drop-Out Mechanism in Repeated-Measures Studies.* Journal of the American Statistical Association, 1995. **90**(431): p. 1112-1121.
23. Roderick, J.A.L. and Y. Wang, *Pattern-Mixture Models for Multivariate Incomplete Data with Covariates.* Biometrics, 1996. **52**(1): p. 98-111.