Trends in iMSMS data - lifestyle factors and MS types

Alpha Diversity – Measurement of complexity of microbes within sample. Relates to richness, evenness, phylogenetic relatedness. Alpha diversity relates to health. Diversity can be good or bad depending on part of body.

Beta Diversity – Compares types of microbes found within 2 samples. Account for relative abundance as well.

Mann Whitney Test on MS and control distributions indicate a significant difference between MS and control in phylum.

Spearman correlation calculates relationship between two variables. +1 meaning positive monotonic(as one variable increases, the other also increases), -1 meaning negative monotonic relationship(as one variable increases, the other decreases), 0 meaning no monotonic relationship.

Bootstrap sampling can also be used to calculate correlation.

Highest correlation in slide 37— 0.48 and -0.43. In conclusion, it does not seem like stronger correlations(beyond 0.5 and -0.5) exist between two variables.

**Possibly read papers on slide 38.**

Screens/Templates Details

It seems like the template for Galileo is already there. The interface(from what can be seen) can be used by users to ask hypothetical questions.

Paper Summary

Docent: Transforming personal intuitions to scientific hypotheses through content learning and process training: **This paper’s idea is eerily similar to ours. What is the difference?** Specification of MS perhaps? Perhaps this paper is more general? **Perhaps read this paper. Read all the papers that are in this file.**

Lessons learned, challenges & feedback

Heatmaps and scatterplots help visualize some relationships but don’t solve all the problems due to data’s complexity. Confounding factors make it difficult to answer specific hypothesis questions. Large sample size and many variables make it hard to isolate clear relationships. It's essential to connect with MS experts to understand the questions they want answered & how tools can do this. Reproducing MS and microbiome relations can lead to more reliable results.

iMSMS Visualization

**Read Papers**

Some things from reproducibility folder

**Papers**

Ideas

Perhaps users have hypothesis that is not present in the data? Perhaps that’s where contribution comes from? Visualization will be a way to check whether the hypothesis is in the data. Will hypothesis include questions about microbial diversity? Is that something that MS people are interested in? Catherine mentions that Chao1 and Shannon may not be conducive to the average person. Does the average MS person use reddit vigorously? Can we make a potential benefit out of MS patients not having to invoke insurance if their hypotheticals are answered/assuring? Can we have potential outcomes from the visual design?

Read the design from TummyTrials. How do participants look at the results of their self experiments?

Read paper that Vineet sent in email