























## Machine Learning framework

- Based on scikit-learn Python package
- Using Random Forest classification and regression algorithms
- 100 bootstrap iterations with 80/20 random split of training and testing folds
- Remove any twin from the training fold if the twin-pair present in the testing fold
- PREDICT 1 US used as independent validation to test the models trained on PREDICT 1 UK
- Microbiome data based on the MetaPhIAn 3 species relative abundances

March 22 2021











## What are the microbes most associated with foods?

Species clustering according to "healthy / less healthy" animal-based foods was distinct and overlapping with those for "healthy / less healthy" plant-based foods

More healthy animal-based foods were associated with a healthier cluster, suggesting that healthier animal-based foods are associated with a more favorable diet-microbiome signature, although this may reflect an overall healthier dietary pattern



arch, 23 2021



















## Gut transit time measured via blue poo is strongly linked with microbiome composition in PREDICT 1

- PREDICT 1 participants were given 2 blue muffins and then selfreported the first appearance of blue poo
- We found that gut transit time is more linked with the microbiome than traditional measures (like Bristol Stool Form scale and number of bowel movements)
- This can be used to measure gut transit time in very large-scale cohorts

23 202



22

