Special Issue: The Gut Microbiome and Aging

Comparing Analytical Methods for the Gut Microbiome and Aging: Gut Microbial Communities and Body Weight in the Osteoporotic Fractures in Men (MrOS) Study

Michelle Shardell, PhD,^{1,*} Neeta Parimi, MS,² Lisa Langsetmo, PhD,³ Toshiko Tanaka, PhD,⁴ Lingjing Jiang, MS,⁵ Eric Orwoll, MD,⁶ James M. Shikany, DrPH,⁷ Deborah M. Kado, MD, MS,⁵ and Peggy M. Cawthon, PhD^{2,8}; for the Osteoporotic Fractures in Men (MrOS) Study Group

¹Department of Epidemiology, Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland. ²Research Institute, California Pacific Medical Center, San Francisco. ³Division of Epidemiology and Community Health, University of Minnesota, Minneapolis. ⁴Translational Gerontology Branch, National Institute on Aging Intramural Research Program, Baltimore, Maryland. ⁵Departments of Family Medicine and Public Health and Internal Medicine, University of California, San Diego School of Medicine, La Jolla. ⁶Department of Medicine, Oregon Health & Sciences University, Portland. ⁷Division of Preventive Medicine, School of Medicine, University of Alabama at Birmingham, Birmingham. ⁸Department of Epidemiology and Biostatistics, University of California, San Francisco.

*Address correspondence to: Michelle Shardell, PhD, Department of Epidemiology and Public Health, Institute for Genome Sciences, 670 W Baltimore Street, Baltimore, MD 21201. E-mail: mshardell@som.umaryland.edu

Received: August 30, 2019; Editorial Decision Date: January 21, 2020

Decision Editor: Michal Masternak, PhD

Abstract

Determining the role of gut microbial communities in aging-related phenotypes, including weight loss, is an emerging gerontology research priority. Gut microbiome datasets comprise relative abundances of microbial taxa that necessarily sum to 1; analysis ignoring this feature may produce misleading results. Using data from the Osteoporotic Fractures in Men (MrOS) study ($n = 530$; mean [SD] age = 84.3 [4.1] years), we assessed 163 genera from stool samples and body weight. We compared conventional analysis, which does not address the sum-to-1 constraint, to compositional analysis, which does. Specifically, we compared elastic net regression (for variable selection) and conventional Bayesian linear regression (BLR) and network analysis to compositional BLR and network analysis; adjusting for past weight, height, and other covariates. Conventional BLR identified *Roseburia* and *Dialister* (higher weight) and *Coprococcus-1* (lower weight) after multiple comparisons adjustment ($p < .0125$); plus *Sutterella* and *Ruminococcus-1* ($p < .05$). No conventional network module was associated with weight. Using compositional BLR, *Coprococcus-2* and *Acidaminococcus* were most strongly associated with higher adjusted weight; *Coprococcus-1* and *Ruminococcus-1* were most strongly associated with lower adjusted weight ($p < .05$), but nonsignificant after multiple comparisons adjustment. Two compositional network modules with respective hub taxa *Blautia* and *Faecalibacterium* were associated with adjusted weight ($p < .01$). Findings depended on analytical workflow. Compositional analysis is advocated to appropriately handle the sum-to-1 constraint.

Keywords: Compositional analysis, Frailty, Bayesian regression, Network analysis

High-throughput sequencing technology has accelerated research of the human gut microbiome. Researchers are interested in identifying aging-related causes and health consequences of gut microbiota in order to develop healthspan-enhancing treatments. One potential consequence of gut dysbiosis is weight loss in old age. Weight loss, or “shrinking,” is a frailty syndrome criterion associated with disability

and mortality (1) that may reflect underlying health conditions (2). Mice and human studies suggest that gut microbiota may affect body weight by regulating appetite and energy absorption (3–6). Gut microbiota and obesity have been assessed in younger adults (7), but no study has examined gut microbiota and body weight in older adults emphasizing low weight.

One barrier to microbiome and aging research is the complex structure of gut microbiota data. Gut microbiome profiles are determined by sequencing the 16S ribosomal RNA (rRNA) gene and classifying the taxon of each rRNA “read.” Therefore, data for each sample is a vector of taxa read counts. Salient features of microbiome data include (i) high dimensionality due to many taxa; (ii) sparsity, where many taxa have zero counts; and (iii) compositionality (8), meaning that since total read counts differ by sample and are constrained by the DNA sequencer, taxon read count does not reflect absolute microbial abundance, only relative abundance; thus, counts are often normalized by a sample’s total read count and are constrained to sum to 1. Additional challenges of gut microbiome research in older adults, as compared with younger adults, are the differences in interpretation (shrinking vs obesity) and differences in confounders (eg, multiple medications vs work-related stress).

Previous work has compared common analytical approaches for microbiota *outcomes* and has found that conclusions are sensitive to approach (9), mainly due to different assumptions about the microbiota distribution. However, no work has compared approaches for microbiota *exposures*, where the microbiota distribution may be unspecified, but the sum-to-1 constraint still must be addressed. Microbiome researchers have stressed the importance of addressing the constraint (8); although others have suggested that it can be ignored in high-diversity settings like the gut (10).

To fill important scientific and analytical gaps in gut microbiome and aging research, we assess associations of gut microbiota with body weight in older men enrolled in the Osteoporotic Fractures in Men (MrOS) study. We compare a *conventional* analytical workflow that ignores the sum-to-1 constraint to a *compositional* workflow that explicitly addresses the constraint while adjusting for confounders relevant for older adults.

Methods

Study Sample

Participants in this study are a subsample of community-dwelling older men enrolled in MrOS. MrOS is a cohort study of 5,994 men aged ≥ 65 years recruited from six U.S. academic medical centers (University of Alabama at Birmingham [Birmingham, AL]; University of Minnesota [Minneapolis, MN]; Stanford University [Palo Alto, CA]; University of Pittsburgh [Pittsburgh, PA]; Oregon Health and Science University [Portland, OR]; and University of California, San Diego [San Diego, CA]) between 2000 and 2002 (11). Eligibility criteria included ability to walk unassisted and no bilateral hip replacements. All sites’ Institutional Review Boards approved the study; all participants provided written informed consent.

Participants attended up to four clinic visits through May 2016. Of the 1,841 men who completed the fourth clinic visit between May 2014 and May 2016, 1,328 men were invited to participate in the ancillary microbiome study starting in March 2015. In total, 982 men agreed to provide a stool sample; collection was concurrent with the fourth clinic visit, except for those who already completed the visit who were ineligible. Specimens from the first 599 men were selected for microbiome analysis (12). We excluded 41 men who had taken antibiotics in the past 30 days and 28 men with missing covariate or outcome data, leaving 530 men for analysis.

Stool Collection, Microbiome Assessment, and Data Processing

Sample collection and measurement have been detailed elsewhere (12). Briefly, The OMNIgene-GUT stool/feces collection kit

(OMR-200, DNA Genotek, Ottawa, Canada) was used for home stool sample collection. Participants mailed samples to the Portland site for processing. Specimens were shipped on dry ice to Alkek Center for Metagenomics and Microbiome Research, Baylor College of Medicine in Houston, TX, and stored at -80°C until 16S rRNA gene sequencing. DNA was extracted in the Hamilton STARlet platform following standard MoBio PowerMag Soil DNA (MoBio Laboratories, Inc., Carlsbad, CA) protocol. DNA was subjected to 16Sv4 rDNA amplification using primers 515F and 806R by polymerase chain reaction. Amplicons were pooled and sequenced on an Illumina MiSeq reagent kit v2 (2×250 bp); resulting sequences were demultiplexed based on unique molecular barcodes (13).

Amplicon sequence errors were corrected using the divisive amplicon denoising algorithm (DADA) via R Statistical Software package DADA2 v1.7.0 (14) with default options unless noted. DADA infers amplicon sequence variants without constructing operational taxonomic units. Briefly, demultiplexed fastq files were filtered and trimmed to lengths of 240 in forward and 200 in reverse, with maximum expected errors of 1 and 2, respectively. Error rates were estimated using a minimum 10^6 reads. Sequence variants were inferred by dereplication, denoising, merging paired reads, and sequence tabulation. Next, chimera were recognized and removed. Taxonomy assignment was carried out by mapping denoised sequences to SILVA Database v128. Phylogeny was constructed by aligning sequences using the msa package with method Clustal Omega (15) then calling FastTree v2.1.9 (16). The final microbiome dataset comprised 18,083,938 reads. In total, 10,201 unique taxa were identified and agglomerated into 388 genera. After filtering genera with $<5\%$ prevalence and $<0.0001\%$ overall relative abundance, the final taxonomy comprised 163 genera, 52 families, 32 orders, 26 classes, and 12 phyla.

Outcome and Covariate Assessment

The study outcome was measured body weight at visit 4 (kg). Rather than normalize by dividing by height² to compute body mass index (BMI), we normalized by weight and height measured at visit 3 (between March 2007 and March 2009) via statistical adjustment. This approach overcomes limitations of BMI (17), allows interpretation of regression coefficients as both adjusted mean weight differences and differences in changes (see statistical analysis, [Supplementary Methods](#) and ref. (18)), and disentangles weight from measured height, which depends on posture. Additional covariates included study site; 16S read counts; demographics (age, race/ethnicity [non-Hispanic white, non-Hispanic black, other]); high-school graduate (yes, no); smoking status (never, former, current); self-rated health (4-point Likert scale from “poor” to “excellent”); self-reported physical activity (Physical Activity Scale for the Elderly [PASE]) (19); self-reported history of diseases (cancer, rheumatoid arthritis); and medications assessed by inventory (antidepressives, β -blockers, proton pump inhibitors, calcium-channel blockers, oral corticosteroids, diabetes medications). Selected diseases or medications may influence weight and gut microbiota directly or indirectly due to medications used for treatment (20). We included diet measures assessed by food frequency questionnaire: probiotic use $\geq$ twice/month (fermented cabbage, milk, or tea; dark chocolate; miso soup; pickles; probiotic drinks; tempeh; yogurt); protein (g/day); saturated fat (g/day); unsaturated fat (g/day); fiber (g/day); other carbohydrate (g/day); and alcohol (drinks/day); and scores for two diet patterns derived from factor analysis, Western diet pattern (eg, red meat, fried food, fat) and prudent diet pattern (eg, fruit, vegetables, whole grains, lean meat) (21). The relative timing of measures resembled that of

a recent gut microbiome and weight change study of middle-aged adults (7). See [Supplementary Methods](#) for conceptual framework.

Statistical Analysis

We computed descriptive statistics by quartiles of weight. We compared categorical variables using chi-square or Fisher's Exact Test; and continuous variables using analysis of variance. We also compared men who completed the fourth clinic visit by whether they were selected to participate in the microbiome substudy.

Our general statistical approach was motivated by the conceptual framework in the [Supplementary Methods](#). We regressed weight on microbiota metrics, covariates assessed at visit 4, weight and height measured at visit 3, study site, and demographics described above. Since we adjusted for visit 3 weight, inference is equivalent to that with weight change as the outcome (18) ([Supplementary Methods](#)). This general approach was carried out using two analytical workflows, a *conventional* analysis and a *compositional* analysis, to relate microbiota with covariate-adjusted weight. A sensitivity analysis was additionally performed on a subsample that did not report taking proton pump inhibitors. Total DNA reads were constrained by the sequencer; thus, higher observed count of one taxon implies lower observed count of at least one other taxon. When taxon count is normalized by total read count, all samples necessarily sum to 1. Thus, data are compositional (22).

We measured within-person microbial diversity (alpha-diversity) of unfiltered data (388 genera) using the Shannon index, a common measure appropriate for compositional data (8). We linearly regressed weight on alpha-diversity and covariates. Next, we computed between-person diversity (beta-diversity) of unfiltered data using weighted UniFrac distance, which uses phylogenetic information (23), but does not address the sum-to-1 constraint. Weighted UniFrac is sensitive to uneven read depth, and rarefying data has been shown to perform better than other methods (24). Therefore, despite data loss (25), we rarefied data to equal depth only for weighted UniFrac. Rarefied counts were log-transformed and centered to be non-negative. We also measured beta-diversity using Aitchison distance for compositional data (22). Aitchison distance is a Euclidean distance of centered log ratios—log-ratios of taxon relative abundance to geometric mean relative abundance ([Supplementary Methods](#)). We performed principal coordinate analysis with both measures. MiRKAT, a kernel machine regression approach (26), tested the multivariate association of microbial composition with covariate-adjusted weight using both measures.

We assessed individual microbial taxa and weight using both a conventional and compositional approach in which microbiota were the exposures. Microbial taxa are interrelated, thus isolating the association of one taxon with weight requires adjusting for other taxa; however, microbiota relative abundance data are high-dimensional and perfectly collinear due to the sum-to-1 constraint. Conventional analysis selected taxa using elastic net regression (ENR) of weight on taxa relative abundance (27), followed by Bayesian linear regression (BLR) of weight on relative abundance of selected taxa and covariates to obtain de-biased coefficients. We assessed a range of values for the ENR sparsity parameter and used cross validation to select the final value. BLR coefficients for each selected taxon are interpreted as mean weight difference per 1% higher relative abundance adjusting for other selected taxa and covariates. Thus, interpretation depends on selected taxa. We performed compositional BLR by transforming taxa relative abundances into isometric log ratios (ILRs). ILR measures a taxon's *relative dominance* by comparing log taxon relative abundance to log geometric mean relative

abundances of remaining taxa, holding relative dominances between remaining taxa constant ([Supplementary Methods](#)) (28). Transformed BLR coefficients are interpreted as mean weight difference per doubling of a taxon's relative dominance adjusting for relative dominances of remaining taxa and covariates. Weak prior distributions in BLR addressed high dimensionality (29). Multiple imputation of 50 data sets using Dirichlet distributions and applying Rubin's rules addressed zeros for logarithms (30) by treating zeros as values below a detection limit. We adjusted individual taxon *p* values by comparing *p* values to .05 divided by the number of principal components (the effective number of independent tests) determined from a scree plot. This approach, originally proposed for genetic studies, overcomes overly conservative Bonferroni corrections for correlated tests, performs well at controlling type-I error (31), and unlike false discovery rate, ensures the same criterion for both workflows. Heatmaps of top taxa across weight quartiles were generated by arithmetic means with conventional hierarchical clustering and by geometric means with clustering by Aitchison distance.

Lastly, due to taxa inter-relationships, we performed a network analysis to identify taxa modules and to determine associations of modules with covariate-adjusted weight. This analysis involved three steps. First, we computed correlations between taxa. We used Pearson correlations (known to be biased, due to the sum-to-1 constraint) of relative abundance for conventional analysis, and Sparse Correlations for Compositional Data (SparCC) (32), which handles the constraint ([Supplementary Methods](#)). Second, we used the two correlation matrices to perform weighted gene coexpression network analysis (WGCNA) (33) to identify taxa modules. We specified an unsigned network so that modules may include both negatively and positively correlated taxa. We specified a minimum module size of 7 (average number of genera per class in the data set). We identified each module's hub taxon, defined as the taxon with the highest intramodule connectivity (33) ([Supplementary Methods](#)). Third, principal components analysis computed each module's "eigntaxon," analogous to an eigengene (33), defined as the first principal component, and regressed weight on eigntaxa and covariates using BLR. Eigntaxa were computed using both conventional and compositional correlations ([Supplementary Methods](#)).

All analyses were performed using R Statistical Software version 3.5.1. $p < .05$ ($p < .05/\#$ of principal components for individual taxa) was considered statistically significant. We used packages *MiRKAT* (26), *glmnet* (27), *arm* (29), *WGCNA* (33), *phyloseq* (34), *vegan* (35), *compositions* (36), *robCompositions* (37), and *SparCC* in *SpiecEasi* (38). These packages enable a complete workflow using R. Compositional R code is in [Supplementary Methods](#).

Results

Participants had mean (standard deviation [SD]) time of 7.4 (0.4) years between visits 3 and 4 and mean (SD) age of 84.3 (4.1) years at visit 4, which was negatively associated with weight ($p < .001$; [Table 1](#)). Men with the lowest body weight were least likely to be white, had the lowest visit 3 weight and height, and lowest saturated fat intake and Western diet scores; men with the highest weight had the lowest proportion of high school graduates (all $p < .05$). Men selected for the microbiome substudy were less likely to be white, nonsmokers, or use proton pump inhibitors and consumed less unsaturated fat and had lower Western diet scores, but they were more likely to visit the Portland and San Diego study sites, were more physically active, and had a longer time between visits than men who were not selected ($p < .05$; [Supplementary Table](#)

Table 1. Characteristics of 530 Participants from the MrOS Study by Quartile of Visit 4 Weight

Characteristics	Q1: ≤ 71 kg (N = 139)	Q2: 71 to 79 kg (N = 138)	Q3 79 to 87 kg (N = 121)	Q4 > 87 kg (N = 132)	p Value
	Mean (SD) or n (%)	Mean (SD) or n (%)	Mean (SD) or n (%)	Mean (SD) or n (%)	
Age (years)	85.3 (4.5)	84.6 (4.1)	83.8 (3.5)	83.3 (3.8)	<.001
Race:					<.001
Black	8 (5.8)	3 (2.2)	5 (4.1)	2 (1.5)	
White	105 (75.5)	126 (91.3)	112 (92.6)	123 (93.2)	
Other Race	26 (18.7)	9 (6.5)	4 (3.3)	7 (5.3)	
High-School Graduate	133 (95.7)	135 (97.8)	121 (100.0)	125 (94.7)	.037
Visit 3 Weight (kg)	70.1 (6.7)	78.8 (4.8)	85.9 (5.1)	98.0 (10.1)	<.001
Visit 3 Height (cm)	169.8 (6.1)	173.1 (5.8)	175.6 (5.5)	178.2 (6.6)	<.001
Physical Activity Scale for the Elderly	128.4 (66.8)	131.8 (72.3)	118.5 (60.4)	112.9 (65.6)	.076
Smoking Status:					.17
Current Smoker	5 (3.6)	1 (0.7)	0 (0.0)	3 (2.3)	
Former Smoker	60 (43.2)	59 (42.8)	64 (52.9)	63 (47.7)	
Nonsmoker	74 (53.2)	78 (56.5)	57 (47.1)	66 (50.0)	
Alcohol (drinks/d)	0.24 (0.24)	0.25 (0.24)	0.28 (0.25)	0.23 (0.24)	.32
Good or Excellent Self-Rate Health	126 (90.6)	128 (92.8)	107 (88.4)	113 (85.6)	.26
Protein (g/d)	59.4 (24.6)	58.3 (23.1)	63.4 (27.9)	62.7 (27.6)	.30
Saturated Fat (g/d)	18.9 (8.7)	19.8 (8.7)	21.8 (10.0)	23.2 (10.5)	.001
Unsaturated Fat (g/d)	43.9 (21.2)	47.3 (24.3)	49.1 (22.2)	50.9 (24.2)	.080
Fiber (g/d)	17.0 (8.4)	15.5 (8.0)	15.7 (7.9)	15.3 (7.7)	.29
Nonfiber Carbohydrate (g/d)	154.3 (58.7)	153.4 (64.7)	156.0 (69.3)	157.0 (67.6)	.97
Western Diet Pattern	-0.37 (0.98)	-0.11 (0.86)	0.07 (1.09)	0.20 (1.05)	<.001
Prudent Diet Pattern	0.17 (1.09)	-0.05 (1.01)	0.00 (1.01)	-0.11 (1.09)	.15
Probiotic Food Intake	117 (84.2)	110 (79.7)	103 (85.1)	110 (83.3)	.66
Cancer	56 (40.3)	62 (44.9)	67 (55.4)	70 (53.0)	.05
Rheumatoid Arthritis	10 (7.2)	10 (7.2)	8 (6.6)	14 (10.6)	.62
Antidepressives	7 (5.0)	15 (10.9)	15 (12.4)	9 (6.8)	.12
β -blockers	37 (26.6)	48 (34.8)	48 (39.7)	48 (36.4)	.14
Proton Pump Inhibitor	25 (18.0)	31 (22.5)	31 (25.6)	19 (14.4)	.12
Calcium Channel Blockers	23 (16.5)	26 (18.8)	26 (21.5)	24 (18.2)	.78
Oral Corticosteroids	2 (1.4)	5 (3.6)	4 (3.3)	2 (1.5)	.57
Diabetes Medication	16 (11.5)	14 (10.1)	12 (9.9)	23 (17.4)	.21
Years Between Visits	7.41 (0.42)	7.45 (0.41)	7.41 (0.45)	7.43 (0.41)	.86
Shannon Diversity	2.59 (0.61)	2.57 (0.59)	2.62 (0.55)	2.55 (0.55)	.80

S1 in the [Supplementary Material](#)). Using conventional arithmetic means, the five most abundant taxa were *Bacteroides* (30.8%), *Faecalibacterium* (7.5%), *Alistipes* (4.5%), *Prevotella*-9 (3.5%), and *Akkermansia* (3.3%). The highest compositional (geometric) means were *Bacteroides* (48.7%), *Faecalibacterium* (7.3%), *Alistipes* (5.2%), a genus in family *Lachnospiraceae* (3.9%), and *Roseburia* (3.8%; [Supplementary Table S2](#) in the [Supplementary Material](#)).

After covariate adjustment, participants with higher Shannon (alpha) diversity tended to weigh more, but the result was not statistically significant (beta coefficient = 0.42, 95% confidence interval [CI], -0.46, 1.30, $p = .35$). Results were similar after excluding participants who reported taking proton pump inhibitors (beta coefficient = 0.39, 95% CI, -0.58, 1.36, $p = .43$). Beta-diversity measured by weighted UniFrac was significantly associated with covariate-adjusted weight ($p = .049$); however, evidence was stronger using Aitchison distance ($p = .004$). The highest-weight men tended to have lower second principal coordinate values derived from UniFrac, whereas the lowest-weight men tended to have lower second principal coordinate values derived from Aitchison distance ([Figure 1](#)). Evidence was weaker after excluding participants who reported taking proton pump inhibitors (UniFrac: $p = .088$; Aitchison: $p = .011$; [Supplementary Figure 1](#)).

Compositional principal component analysis identified four components ([Supplementary Figure 2](#)), thus statistical significance

for individual taxon was $p < .05/4 = .0125$. ENR selected 35 taxa. Nine of the top 10 most significant taxa identified by ENR with BLR were in phylum Firmicutes ([Figure 2](#), [Supplementary Table S3](#)). *Roseburia* and *Dialister* were significantly associated with higher adjusted weight ($p < .005$); men with higher *Sutterella* (phylum Proteobacteria) relative abundance also tended to have higher adjusted weight ($p = .026$). *Coprococcus*-1 was significantly associated with lower weight ($p = .011$). Men with higher *Ruminococcus*-1 also tended to have higher adjusted weight ($p = .033$). Similarly, 9 of the top 10 most significant taxa identified by compositional BLR were in phylum Firmicutes ([Figure 2](#), [Supplementary Table S4](#)). No taxa were significantly associated with weight after multiple comparisons adjustment; however, men with higher relative dominance of *Coprococcus*-2 and *Acidaminococcus* tended to have higher adjusted weight ($p < .05$), and men with higher relative dominance of *Coprococcus*-1 and *Ruminococcus*-1 tended to have lower adjusted weight ($p < .05$). The approaches shared five genera in their respective top 10 (*Roseburia*, *Coprococcus*-1, *Sutterella*, *Ruminococcus*-1, and *Butyrivibrio*). Heatmaps of top 30 taxa from both models (50 unique taxa) using conventional arithmetic mean and compositional mean revealed, respectively, clustering based on overall relative abundance and by relative patterns across weight quartiles ([Figure 3](#)). After excluding $n = 106$ men who reported taking proton pump

Figure 1. Beta-diversity and weight principal coordinate analysis. Beta-Diversity measured using Weighted UniFrac (top) and Aitchison distance (bottom). p value computed using MIRKAT adjusted for covariates. Full color version is available within the online issue.

inhibitors, evidence for association tended to be weaker (larger P), and some of the top 10 taxa differed for both analytical approaches; however, the most significant taxa from the original analyses tended to remain in the top 10 (Supplementary Figure 3).

Network analysis using Pearson correlations identified three modules comprising 8, 7, and 7 taxa, respectively (Figure 4, Supplementary Table S5, Supplementary Figure 4). Modules 1 and 2 had respective hub taxa *Collinsella* (phylum Actinobacteria) and *Ruminococcaceae-UCG-010* (phylum Firmicutes); module eigentaxa were not significantly associated with adjusted weight ($p > .80$) (Supplementary Table S6). Module 3 had hub taxon *Anaerofilum* (phylum Firmicutes), comprised positively correlated taxa from five families in phyla Firmicutes and Actinobacteria (Supplementary Table S5), and its eigentaxon was not quite significantly associated with lower adjusted weight ($p = .088$) (Supplementary Table S6). Findings on module associations with weight differed little after excluding $n = 106$ men who reported

Figure 2. Association of individual taxa with covariate-adjusted weight. Volcano plots from elastic net regression with conventional Bayesian linear regression of relative abundance adjusted for covariates (top) and compositional Bayesian linear regression of isometric log-ratios of relative abundance adjusted for covariates (bottom). Labelled taxa are the top 10 most significant taxa from each respective analysis. Taxa above the upper horizontal line have $p < .05/4 = .0125$; taxa above the lower horizontal line have $p < .05$. Full color version is available within the online issue.

taking proton pump inhibitors (Supplementary Table S6). In contrast, network analysis using SparCC, identified five modules comprising 27, 27, 22, 20, and 9 taxa, respectively (Figure 4, Supplementary Table S7, Supplementary Figure 5). Module 1 had *Blautia* (phylum Firmicutes) as its hub. Its eigentaxon was a ratio of genera in families *Erysipelotrichaceae* and *Lachnospiraceae* to *Coprococcus-2* (also in family *Lachnospiraceae*), a genus in family *Peptococcaceae* and genera in family *Ruminococcaceae* including *Ruminococcaceae-UCG-003*; the eigentaxon was associated with lower adjusted weight ($p < .001$; Supplementary Table S8). Module 2 had hub taxon *Ruminococcaceae-UCG-005* (phylum Firmicutes) and an eigentaxon interpreted as a ratio of genera within family *Ruminococcaceae* to taxa from 5 phyla. Higher eigentaxon scores were not quite significantly associated with lower adjusted weight ($p = .055$; Supplementary Table S8).

Figure 3. Heat maps of mean relative abundance of 50 unique taxa comprising top 30 taxa from individual taxa analysis by weight quartile. Arithmetic means clustered using Euclidean distance (top) and compositional means clustered using Aitchison distance (bottom). Full color version is available within the online issue.

Module 3 had *Faecalibacterium* (phylum Firmicutes) as its hub. Its eigentaxon was a ratio of *Erysipelotrichaceae_UCG-003*, genera from families *Lachnospiraceae* (eg, *Roseburia*) and *Ruminococcaceae*

Modules and Hub Taxa from Pearson WGCNA

- Module 1 (*Collinsella_g*)
 - Module 2 (*Ruminococcaceae_UCG-010_g*)
 - Module 3 (*Anaerofilum_g*)
- Positive
— Negative

Modules and Hub Taxa from SparCC WGCNA

- Module 1 (*Blautia_g*)
 - Module 2 (*Ruminococcaceae_UCG-005_g*)
 - Module 3 (*Faecalibacterium_g*)
 - Module 4 (*Oscillibacter_g*)
 - Module 5 (*Intestinibacter_g*)
- Positive
— Negative

Figure 4. Taxa modules with respective hub taxa derived from weighted gene coexpression network analysis. Modules derived using Pearson correlations (top) and SparCC (bottom). See [Supplementary Tables 4 and 6](#) for taxa and eigentaxon loadings. Edge thickness denotes strength of correlation. Edges with correlation > |0.2| shown. Full color version is available within the online issue.

(eg, *Butyricoccus*), to genera from four phyla including four genera within phylum *Bacteroidetes* (eg, *Butyricimonas*); the eigentaxon was associated with higher adjusted weight ($p = .004$) ([Supplementary Table S8](#)). Modules 4 and 5 had respective hub taxa *Oscillibacter* and *Intestinibacter* in phylum Firmicutes; neither eigentaxon was significantly associated with adjusted weight ($p > .50$; [Supplementary Table S8](#)). Findings on module associations with weight differed little after excluding $n = 106$ men who reported taking proton pump inhibitors ([Supplementary Table S8](#)).

Discussion

Overall, we found evidence that the gut microbiome is associated with covariate-adjusted body weight among community-dwelling

older men. Although alpha-diversity was not significantly associated with weight, both conventional and compositional workflows found significant associations of microbial distance (beta-diversity) with adjusted weight; however, evidence was stronger using Aitchison distance than weighted UniFrac. This finding is consistent with previous work that linked microbial distance with multiple diet and health measures in older adults (39). Weighted UniFrac required rarefaction to handle uneven read depth, which may have increased measurement error, whereas unevenness was naturally handled by Aitchison distance. Phylogeny-based Aitchison distance has been proposed but has not been found to influence results (8).

Findings regarding individual taxa were sensitive to workflow, but with some agreement. Although no individual taxon was significantly associated with weight after adjustment for multiple comparisons in compositional BLR, both workflows identified *Coprococcus-1* and *Ruminococcus-1* as the top two taxa associated with lower weight. The workflows diverged regarding higher weight; compositional BLR identified *Coprococcus-2* and *Acidaminococcus* as the top two taxa whereas ENR with BLR identified *Roseburia* and *Dialister* as the top two taxa; however, both workflows identified *Sutterella* as the third most significant taxon. The two workflows shared 5 taxa within their respective top 10 and 10 taxa within their respective top 30.

Reasons for different results in individual taxa analysis may be due to scale, adjustment, and handling of zero counts. Coefficients of selected taxa from conventional BLR are on the relative abundance scale and depend on which other taxa are selected. Variable selection created a reference that is an amalgamation of unselected taxa. Therefore, 1% higher relative abundance of a selected taxon corresponds with 1% lower relative abundance of the *unselected* taxa amalgamation. In contrast, coefficients from compositional BLR are on log-ratio scale and are adjusted for relative dominance of all other taxa. Therefore, doubling relative dominance of a taxon corresponds to a proportionately lower relative abundance of all other taxa (Supplementary Methods). An intuitive example is comparing two recipes that differ by changing the amount of one ingredient: increasing one ingredient reduces the proportions of all remaining ingredients, but *ratios* of remaining ingredients to each other are unchanged. A caveat is that we cannot differentiate between a higher amount of one ingredient, lower amounts of other ingredients, or a combination of the two, because absolute quantities are unknown. Since the ILR transformation models all taxa, the interpretation does not depend on omitted reference taxa. Modeling log transformations required imputing zeros. We used the Dirichlet distribution for multiple imputation, which is limited by strong parametric assumptions; however, BLR per se did not require these assumptions. Findings from both workflows are in general more modest than those that compare each taxon to an amalgamation of all other taxa, owing to the extensive adjustment for other taxa. However, both workflows were less vulnerable to unmeasured confounding than the simple pairwise analysis often performed with -omics data.

Network analysis with SparCC identified more and larger modules than did network analysis with Pearson correlation. No Pearson correlation eigentaxon was significantly associated with weight, whereas two SparCC eigentaxa were significantly associated with weight. Higher Module 1 eigentaxon scores, which correspond to relatively lower *Coprococcus-2*, were associated with lower weight, in accordance with individual taxon results. Higher scores of the Module 3 eigentaxon, which included genera from the compositional top 10, were associated with higher weight. Different findings between the two workflows may be

due to biases of Pearson correlations with compositional data, which cannot be remedied with Spearman correlations, but can be remedied using centered log ratios (32). SparCC is an accessible way to adapt downstream analyses, such as network analysis, to compositional data.

Top taxa associated with weight (in either direction), either individually or in a module, include butyrate producers from phylum Firmicutes (4). Butyrate is a short-chain fatty acid (SCFA) end product of fermentation of nondigestible carbohydrates, primarily dietary fiber (4). SCFAs have been associated with better body composition (6); possible mechanisms include regulating appetite and energy expenditure (4,6). Butyrate is an essential metabolite in the colon, is the preferred energy source for colon epithelial cells, influences hormonal and immune signaling pathways, and influences gut pH, which affects nutrient absorption (40). Compositional analysis also had *Acidaminococcus* as a top genus associated with higher weight. These amino acid-fermenting bacteria are involved in amino acid absorption and digestion and were positively associated with BMI in a compositional reanalysis of a cross-sectional study (41).

Strengths of this study include a large well-characterized longitudinal cohort of community-dwelling older men and rigorous comparison of two analytical workflows. Both strengths lend novelty to this study. First, research on gut microbiota and weight has focused on obesity in younger adults (7) or disease-specific weight loss (eg, cancer) (42). Very little work has considered microbiome and weight in community-dwelling older adults (43). In this context, careful consideration of the interpretation of results and inclusion of confounding factors that may influence shrinking and the microbiome (eg, medications, health-related events) is needed. Multiple statistical software for microbiome *outcome* data were compared in a published sensitivity analysis (9) and addressed in prominent reviews (8,10). However, this study filled the gap of comparing methods for compositional microbiome *exposure* data.

Despite study strengths, some limitations are worth noting. First, there is potential for selection bias and truncation by death given that data were collected at a visit over a decade after recruitment. Thus, findings may generalize only to the healthiest participants. Indeed, in agreement with previous work (12), participants selected for the microbiome substudy differed from those who were not selected on some factors; thus, adjusting for covariates was particularly important. Second, microbiota relative abundance data are prone to measurement error and limits of detection due to sampling of reads and uneven read counts; however, we addressed these issues by rigorous analysis that included multiple imputation. Alternatives to the Dirichlet distribution used here for zero imputation and other normalization methods should be assessed in future microbiome exposure research. Third, taxon classification is imperfect; however, this affected both analytical workflows. Lastly, as in any observational study where exposure and outcome are concurrently assessed, it is difficult to determine whether the putative exposure (gut microbial profile) influences the putative outcome (weight) or vice versa, or whether confounders were unmeasured. Thus, the role of microbiota is ambiguous here. We mitigated these issues by adjusting for past weight, height, diet, behavior, and other health-related covariates, which may also reduce the influence of selection bias. Not a limitation per se, but a notable challenge of gut microbiome research in older adults is the high frequency of multiple medication use. We addressed this challenge by excluding participants with recent antibiotic use and adjusting for medications and health conditions often requiring medications.

In a sensitivity analysis, we additionally excluded participants who reported taking proton pump inhibitors; this tended to produce weaker results, but a plausible reason for this is the 20% drop in sample size.

In summary, we found robust evidence linking overall gut microbial profiles to body weight, and evidence was stronger using compositional analysis. Findings for individual taxa were stronger using conventional analysis, plausibly due to differences in adjustment for other taxa. Lastly, findings from network analysis were stronger using compositional analysis. This finding is likely due to bias from conventional correlation. These findings are consistent with an overall association between gut microbiota and weight in older men that cannot be reduced to simple profiles. We demonstrated that workflows for high-dimensional -omics data can be adapted to handle compositional microbiome data. More sensitivity analyses of zero handling and normalization are needed to optimize compositional workflows of microbiome exposure data. Compositional analysis is advocated to appropriately handle the sum-to-1 constraint.

Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences* online.

Funding

This work was supported by the National Institutes of Health (grant numbers R01 AG048069, R01 AI116799). MrOS was supported by the National Institutes of Health (grant numbers U01 AG027810, U01 AG042124, U01 AG042139, U01 AG042140, U01 AG042143, U01AG042145, U01 AG042168, U01 AR066160, and UL1 TR000128).

Author Contributions

M.S., L.L., E.O., and P.M.C. were responsible for conception and study design. N.P., E.O., J.M.S., D.M.K., and P.M.C. contributed to data acquisition and management. M.S. performed statistical analysis and drafted the manuscript. All authors contributed to data interpretation and editing the manuscript for content; all authors reviewed the manuscript and approved the final version.

Conflict of Interest

M.S., E.O., J.M.S., D.M.K., and P.M.C. are supported by grants from the National Institutes of Health. L.L. received funding from Abbott Nutrition. Remaining authors have no conflicts of interest.

References

- Strandberg TE, Stenholm S, Strandberg AY, Salomaa VV, Pitkälä KH, Tilvis RS. The “obesity paradox,” frailty, disability, and mortality in older men: a prospective, longitudinal cohort study. *Am J Epidemiol*. 2013;178:1452–1460. doi:10.1093/aje/kwt157
- Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci*. 2001;56:M146–156. doi:10.1093/gerona/56.3.m146
- Turnbaugh PJ, Ley RE, Mahowald MA, et al. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*. 2006;444:1027–1031. doi:10.1038/nature05414
- Frost G, Sleeth ML, Sahuri-Arisoylu M, et al. The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism. *Nat Commun*. 2014;5:3611. doi:10.1038/ncomms4611
- Turnbaugh PJ, Hamady M, Yatsunenko T, et al. A core gut microbiome in obese and lean twins. *Nature*. 2009;457:480–484. doi:10.1038/nature07540
- Byrne CS, Chambers ES, Morrison DJ, Frost G. The role of short chain fatty acids in appetite regulation and energy homeostasis. *Int J Obes (Lond)*. 2015;39:1331–1338. doi:10.1038/ijo.2015.84
- Menni C, Jackson MA, Pallister T, et al. Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain. *Int J Obes (Lond)*. 2017;41:1099–1105. doi:10.1038/ijo.2017.66
- Gloor GB, Macklaim JM, Pawlowsky-Glahn V, et al. Microbiome datasets are compositional: and this is not optional. *Front Microbiol*. 2017;8:2224. doi:10.3389/fmicb.2017.02224
- Zhang X, Nieuwdorp M, Groen AK, Zwinderman AH. Statistical evaluation of diet-microbe associations. *BMC Microbiol*. 2019;19:90. doi:10.1186/s12866-019-1464-0
- Tsilimigras MC, Fodor AA. Compositional data analysis of the microbiome: fundamentals, tools, and challenges. *Ann Epidemiol*. 2016;26:330–335. doi:10.1016/j.annepidem.2016.03.002
- Orwoll E, Blank JB, Barrett-Connor E, et al. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study—a large observational study of the determinants of fracture in older men. *Contemp Clin Trials*. 2005;26:569–585. doi:10.1016/j.cct.2005.05.006
- Abrahamson M, Hooker E, Ajami NJ, Petrosino JF, Orwoll ES. Successful collection of stool samples for microbiome analyses from a large community-based population of elderly men. *Contemp Clin Trials Commun*. 2017;7:158–162. doi:10.1016/j.conctc.2017.07.002
- Caporaso JG, Lauber CL, Walters WA, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J*. 2012;6:1621–1624. doi:10.1038/ismej.2012.8
- Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13:581–583. doi:10.1038/nmeth.3869
- Bodenhofer U, Bonatesta E, Horejs-Kainrath C, Hochreiter S. msa: an R package for multiple sequence alignment. *Bioinformatics*. 2015;31:3997–3999. doi:10.1093/bioinformatics/btv494
- Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One*. 2010;5:e9490. doi:10.1371/journal.pone.0009490
- Tomiya AJ, Hunger JM, Nguyen-Cuu J, Wells C. Misclassification of cardiometabolic health when using body mass index categories in NHANES 2005–2012. *Int J Obes (Lond)*. 2016;40:883–886. doi:10.1038/ijo.2016.17
- Laird N. Further comparative analysis of pretest-posttest research designs. *Am Stat*. 1983;37:329–330. doi:https://doi.org/10.1080/00031305.1983.10483133
- Washburn RA, Ficker JL. Physical Activity Scale for the Elderly (PASE): the relationship with activity measured by a portable accelerometer. *J Sports Med Phys Fitness*. 1999;39:336–340.
- Maier L, Pruteanu M, Kuhn M, et al. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature*. 2018;555:623–628. doi:10.1038/nature25979
- Rogers TS, Harrison S, Judd S, et al. Dietary patterns and longitudinal change in hip bone mineral density among older men. *Osteoporos Int*. 2018;29:1135–1145. doi:10.1007/s00198-018-4388-x
- Aitchison J. *The Statistical Analysis of Compositional Data*. London, UK: Chapman & Hall, Ltd.; 1986. doi:10.1007/978-94-009-4109-0
- Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol*. 2005;71:8228–8235. doi:10.1128/AEM.71.12.8228-8235.2005
- Weiss S, Xu ZZ, Peddada S, et al. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome*. 2017;5:27. doi:10.1186/s40168-017-0237-y
- McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol*. 2014;10:e1003531. doi:10.1371/journal.pcbi.1003531
- Zhao N, Chen J, Carroll IM, et al. Testing in microbiome-profiling studies with mirkat, the microbiome regression-based kernel association test. *Am J Hum Genet*. 2015;96:797–807. doi:10.1016/j.ajhg.2015.04.003
- Friedman JH, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2008;33:1–22. doi:10.18637/jss.v033.i01

28. Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barcelo-Vidal C. Isometric logratio transformations for compositional data analysis. *Math Geol.* 2003;35:279–300. doi:[10.1023/A:1023818214614](https://doi.org/10.1023/A:1023818214614)
29. Gelman A, Hill J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. London, UK: Cambridge University Press; 2006. doi:[10.1017/CBO9780511790942](https://doi.org/10.1017/CBO9780511790942)
30. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ: Wiley; 2004. doi:[10.1002/9780470316696](https://doi.org/10.1002/9780470316696)
31. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol.* 2008;32:361–369. doi:[10.1002/gepi.20310](https://doi.org/10.1002/gepi.20310)
32. Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. *PLoS Comput Biol.* 2012;8:e1002687. doi:[10.1371/journal.pcbi.1002687](https://doi.org/10.1371/journal.pcbi.1002687)
33. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559. doi:[10.1186/1471-2105-9-559](https://doi.org/10.1186/1471-2105-9-559)
34. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One.* 2013;8:e61217. doi:[10.1371/journal.pone.0061217](https://doi.org/10.1371/journal.pone.0061217)
35. Oksanen J, Blanchet FG, Friendly M, et al. vegan: Community Ecology Package. R package version 2.5–5. <https://CRAN.R-project.org/package=vegan>. 2019.
36. van den Boogaart KG, Tolosana-Delgado, R. “compositions”: a unified R package to analyze compositional data. *Comp Geoscience.* 2008;34:320–338.
37. Templ M, Hron K, Filzmoser P. robCompositions: an R-package for robust statistical analysis of compositional data. In: Pawlowsky-Glahn V, Buccianti A, eds. *Compositional Data Analysis: Theory and Applications*. New York, NY: Wiley; 2011. doi:[10.1002/9781119976462](https://doi.org/10.1002/9781119976462)
38. Kurtz ZD, Muller CL, Miraldi ER, et al. Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput Biol.* 2015;11:e1004226. doi:[10.1371/journal.pcbi.1004226](https://doi.org/10.1371/journal.pcbi.1004226)
39. Claesson MJ, Jeffery IB, Conde S, et al. Gut microbiota composition correlates with diet and health in the elderly. *Nature.* 2012;488:178–184. doi:[10.1038/nature11319](https://doi.org/10.1038/nature11319)
40. Martinez RC, Bedani R, Saad SM. Scientific evidence for health effects attributed to the consumption of probiotics and prebiotics: an update for current perspectives and future challenges. *Br J Nutr.* 2015;114:1993–2015. doi:[10.1017/S0007114515003864](https://doi.org/10.1017/S0007114515003864)
41. Shi P, Zhang A, Li H. Regression analysis for microbiome compositional data. *Ann Appl Stat.* 2016;10:1019–1040. doi:[10.1214/16-AOAS928](https://doi.org/10.1214/16-AOAS928)
42. Vivarelli S, Salemi R, Candido S, et al. Gut microbiota and cancer: from pathogenesis to therapy. *Cancers (Basel).* 2019;11. In press. doi:[10.3390/cancers11010038](https://doi.org/10.3390/cancers11010038)
43. Verdi S, Jackson MA, Beaumont M, et al. An investigation into physical frailty as a link between the gut microbiome and cognitive health. *Front Aging Neurosci.* 2018;10:398. doi:[10.3389/fnagi.2018.00398](https://doi.org/10.3389/fnagi.2018.00398)