OPEN ACCESS

ORIGINAL RESEARCH

Longitudinal metabolic and gut bacterial profiling of pregnant women with previous bariatric surgery

Kiana Ashley West ,[1] Chidimma Kanu,[2] Tanya Maric,[2] Julie Anne Kathryn McDonald ,[1] Jeremy K Nicholson,[3] Jia V Li,[1] Mark R Johnson,[2] Elaine Holmes,[1,3] Makrina D Savvidou[2]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2019-319620).

[1]Department of Metabolism, Digestion and Reproduction, Faculty of Medicine, Imperial College London, London, UK
[2]Obstetrics & Gynaecology, Chelsea and Westminster Hospital, Institute of Reproductive Developmental Biology, Department of Metabolism, Digestion & Reproduction, Faculty of Medicine, Imperial College London, London, UK
[3]Australian National Phenome Centre, Health Futures Institute, Murdoch University, Murdoch, Western Australia, Australia

Correspondence to
Elaine Holmes, Department of Metabolism, Digestion and Reproduction, Faculty of Medicine, Imperial College London, London SW7 2AZ, UK; elaine.holmes@imperial.ac.uk

KAW and CK contributed equally.

Received 8 August 2019
Revised 3 December 2019
Accepted 4 December 2019
Published Online First 21 January 2020

© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: West KA, Kanu C, Maric T, *et al*. *Gut* 2020;**69**:1452–1459.

ABSTRACT

Objective Due to the global increase in obesity rates and success of bariatric surgery in weight reduction, an increasing number of women now present pregnant with a previous bariatric procedure. This study investigates the extent of bariatric-associated metabolic and gut microbial alterations during pregnancy and their impact on fetal development.

Design A parallel metabonomic (molecular phenotyping based on proton nuclear magnetic resonance spectroscopy) and gut bacterial (16S ribosomal RNA gene amplicon sequencing) profiling approach was used to determine maternal longitudinal phenotypes associated with malabsorptive/mixed (n=25) or restrictive (n=16) procedures, compared with women with similar early pregnancy body mass index but without bariatric surgery (n=70). Metabolic profiles of offspring at birth were also analysed.

Results Previous malabsorptive, but not restrictive, procedures induced significant changes in maternal metabolic pathways involving branched-chain and aromatic amino acids with decreased circulation of leucine, isoleucine and isobutyrate, increased excretion of microbial-associated metabolites of protein putrefaction (phenylacetlyglutamine, *p*-cresol sulfate, indoxyl sulfate and *p*-hydroxyphenylacetate), and a shift in the gut microbiota. The urinary concentration of phenylacetylglutamine was significantly elevated in malabsorptive patients relative to controls (p=0.001) and was also elevated in urine of neonates born from these mothers (p=0.021). Furthermore, the maternal metabolic changes induced by malabsorptive surgery were associated with reduced maternal insulin resistance and fetal/birth weight.

Conclusion Metabolism is altered in pregnant women with a previous malabsorptive bariatric surgery. These alterations may be beneficial for maternal outcomes, but the effect of elevated levels of phenolic and indolic compounds on fetal and infant health should be investigated further.

INTRODUCTION

Currently more than a third of women are classified as overweight or obese[1] with obesity reaching epidemic levels globally. Obesity is associated with a number of adverse metabolic effects resulting in cardiovascular disease, metabolic syndrome, diabetes and cancer.[2] Bariatric surgery (BS) has proven a successful treatment modality for lasting weight loss and has been shown to reduce obesity-related morbidity and mortality.[3 4] Altered bile flow, reduction in gastric capacity, anatomical gut rearrangement and altered nutrient flow, vagal manipulation and modulation of enteric gut hormones, collectively referred to as the BRAVE effects,[3 5] have been proposed to contribute to the mechanisms of weight loss and diabetes resolution in bariatric patients. There are two main types of BS: the restrictive (RES) procedures such as gastric banding and sleeve gastrectomy, which reduce the stomach size, and the malabsorptive/mixed (MAL) type, such as bilio-pancreatic diversion and Roux-en-Y-gastric bypass, which aim to decrease caloric absorption.[6] BS has been shown to impact metabolic profiles reflecting altered metabolism and changes in the gut microbiome that play a role in improved lipid and glucose metabolism.[7–11] The contribution of the gut mictobiota to host metabolism and

Significance of this study

What is already known on this subject?
- ► Bariatric surgery, particularly the malabsorptive type, alters the gut microbiota and host metabolism.
- ► Pregnancy following bariatric surgery, especially a malabsorptive procedure, is associated with a reduced prevalence of maternal gestational diabetes and increased prevalence of small for gestational age babies.

What are the new findings?
- ► We demonstrated that malabsorptive surgery-induced changes in the maternal gut microbiota persist throughout pregnancy compared with pregnancies with similar maternal body mass index but no history of such surgery, and may have a trans-generational impact.
- ► Gut microbial changes in malabsorptive patients are associated with increased excretion of protein putrefaction metabolites during pregnancy.

How might it impact on clinical practice in the foreseeable future?
- ► The impact of maternal malabsorption on fetal health warrants further investigation to determine the most appropriate type of weight loss surgery for women of reproductive age.

Gut: first published as 10.1136/gutjnl-2019-319620 on 21 January 2020. Downloaded from http://gut.bmj.com/ on April 27, 2024 by guest. Protected by copyright.

BMJ

metabolic control is well-established. Microbes degrade dietary components that are undigested by the host, convert dietary molecules into bioactive metabolites, influence host response to drugs and participate in many homeostatic processes and signalling pathways.[12 13] However, the mechanisms by which BS confers metabolic benefit and the potential role of the microbiome in achieving this are poorly understood.

Evidence continues to accumulate that pregnancy post-BS is associated with lower prevalence of gestational diabetes mellitus (GDM) and large for gestational age neonates but higher risk of small for gestational age neonates and late preterm delivery compared with pregnancies in obese women without a previous BS.[14] To investigate the impact of previous BS on maternal phenotype during pregnancy, we adopted a longitudinal metabolic (serum and urine) and gut bacterial (faeces) profiling strategy using proton nuclear magnetic resonance (^{1}H NMR) spectroscopy and 16S rRNA gene amplicon sequencing, respectively, to characterise women with previous BS (RES and MAL) throughout the course of their pregnancy. Profiles were compared against women with similar early pregnancy body mass index (BMI) and no history of BS (NBS). For a subset of the cohort, we also characterise neonatal cord serum and urine metabolomes to assess potential transgenerational metabolic effects of BS.

METHODS

Study population, sampling and clinical data

The population is part of an ongoing prospective cohort study investigating the impact of maternal BS on perinatal outcomes. All women gave written, informed consent for their data and samples to be used. Pregnant women with (n=47) and without (n=118) previous BS were recruited from May 2015 to April 2017 at Chelsea and Westminster Hospital (London, UK) as previously described.[15] Women were seen at five time points during pregnancy (**T1**: 11^{+0}–14^{+0}, **T2**: 20^{+0}–24^{+0}, **T3**: 28^{+0}–30^{+0}, **T4**: 30^{+0}–33^{+0} and **T5**: 35^{+0}–37^{+6} weeks gestation) and within 72 hours of delivery (**T6**). Maternal blood (serum) and urine samples were collected at each visit while faecal samples were requested at the T1, T2 and T4 visits (see online supplementary figure S1 and table S1). A full oral glucose tolerance test (2 hour, 75 g) was conducted at T3 and maternal insulin resistance was calculated using the homeostatic model assessment for insulin resistance (HOMA-IR=fasting serum insulin (µU/L) × fasting glucose (mmol/L)/22.5). Estimated fetal weight was calculated by trans-abdominal ultrasound scans at T2, T4 and T5. At T6, birth weight was recorded, percentiles for the gestation were calculated and, where possible, neonatal samples (cord serum and urine) were collected. All samples were stored at −80°C for future analysis. For the study population, women with diagnosis of type 2 diabetes mellitus or GDM (due to the effect of diabetes on the metabolic profile) and those that had a miscarriage were excluded. Only NBS participants with a BMI of 25–50 kg/m^2 at T1 were included to match the BMI range of the included bariatric patients at T1.

Metabolic profiling of biofluid samples

Serum and urine samples were prepared according to an established protocol.[16] ^{1}H NMR spectra were acquired on a Bruker 600 MHz spectrometer (Bruker BioSpin) following a published method[16] (see online supplementary methods for further detail). Methodology for preprocessing of spectral data is described in online supplementary methods. Multivariate modelling of the spectral data was performed in the software package SIMCA V.14.1 (Sartorius Stedim Biotech). Principal component analysis (PCA) was used to assess variation in metabolic profiles over all time points and to identify extreme outliers to exclude from supervised models for each time point. Orthogonal partial least squares discriminant analysis (OPLS-DA) was used to identify spectral variables (relating to specific chemical compounds) that contributed to discrimination of clinical classes (pairwise comparisons between NBS, MAL and RES groups) at each time point. From the discriminatory NMR peaks, metabolite identities were confirmed using statistical tools,[17 18] 2D NMR experiments and by spiking in authentic standards (see online supplementary methods). Relative concentrations of each discriminatory metabolite were calculated by integrating a representative peak of that metabolite. Downstream analysis was performed in the R software environment.[19] For each discriminatory metabolite identified in the cross-sectional analysis, time series curves were generated through a spline-fitting method implemented in 'santaR' (see online supplementary methods) to visualise their behaviour in different groups over the pregnancy time course. Individuals with at least five data points were included. Correlation analysis including partial correlation to adjust for confounders is detailed in the online supplementary methods along with a list of all R packages used. Missing data were excluded from calculations. All p values were adjusted (p_{adj}) where necessary to control for the false discovery rate according to the Benjamini-Hochberg method.[20] An alpha of 0.05 was used for p and p_{adj} values.

Gut bacterial community profiling

Stool samples were randomised for processing and DNA was extracted (see online supplementary methods) using the PowerLyzer PowerSoil DNA Isolation Kit (Mo Bio). 16S rRNA gene amplicon sequencing targeting the V1-V2 regions was performed on the Illumina MiSeq platform as previously described.[21] Raw reads were processed in the R software environment[19] following a published workflow[22] which includes amplicon denoising implemented in 'DADA2'.[23] See online supplementary methods for full details. Functions in the 'vegan' R package were used to calculate Shannon Diversity Indices (α-diversity) on data rarefied to the minimum sequencing depth and Bray-Curtis dissimilarity (β-diversity) on log-transformed data (pseudocount of 1 added to each value). Significance of group separation in β-diversity was assessed by permutational multivariate analysis of variance. Changes in relative abundance were tested at each taxonomic rank from phylum to genus using the Mann-Whitney U test while differentially abundant 16S rRNA gene sequences were identified using 'DESeq2'.[24] For 'DESeq2' analysis, data were pooled for each individual rather than analysing distinct time points.

Integrative analysis of metabolic and taxonomic data

Relationships between the serum, urine and faecal data sets were modelled using the DIABLO method in 'mixOmics'.[25] This is a multi-block latent variable-based approach which aims to identify concordance between multiple data sets. Metabolites significant throughout the time course (serum: leucine, isoleucine, isobutyrate, D-β-hydroxybutyrate; urine: phenylacetylglutamine (PAG), *p*-cresol sulfate (PCS), indoxyl sulfate (IS), *p*-hydroxyphenylacetate (PHPA), unknown, α-ketoisovalerate, creatinine) and a subset of bacterial genera (log-transformed; selected using the least absolute shrinkage and selection operator (LASSO) penalisation method implemented in 'mixOmics') were modelled. Sampling points for each individual, where matching microbiome and metabolite data were obtained (T1, T2 and T4), were included in the model.

Gut: first published as 10.1136/gutjnl-2019-319620 on 21 January 2020. Downloaded from http://gut.bmj.com/ on April 27, 2024 by guest. Protected by copyright.

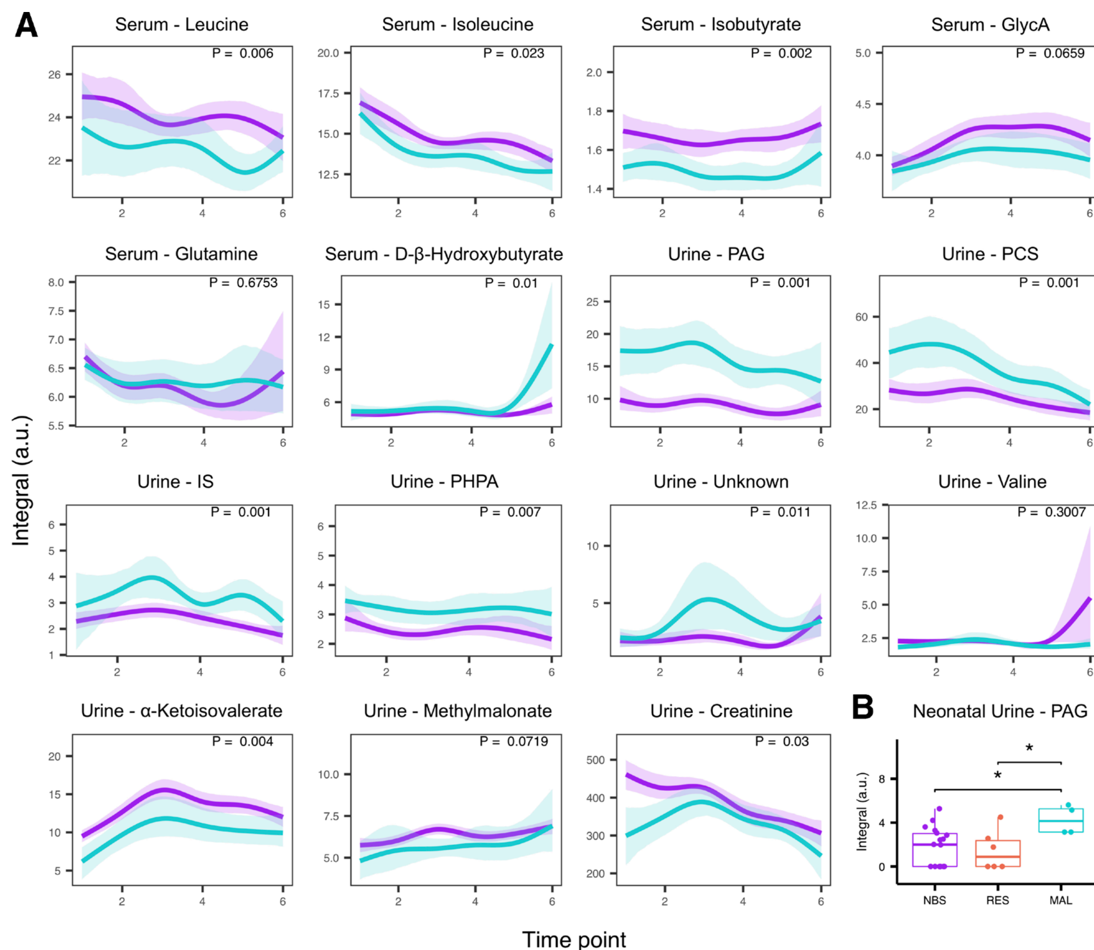

Figure 1 Modelling of metabolite concentrations. (A) Time series analysis of discriminatory metabolites with lines representing group mean curves while shaded areas denote 95% confidence bands. Malabsorptive (MAL, blue) patients relative to No bariatric surgery (NBS, purple) controls. (B) Relative concentration of phenylacetylglutamine (PAG) is increased in neonatal urine from the MAL group compared with the NBS or restrictive (RES) groups. NMR peak integrals are reported in arbitrary units (a.u.). **P*<0.05 (Mann-Whitney U test); GlycA, *N*-acetyl glycoprotein; IS, indoxyl sulfate; PCS, *p*-cresol sulfate; PHPA, *p*-hydroxyphenylacetate.

RESULTS

Characteristics of the study participants

Of the 165 women enrolled in the study, 54 were excluded from longitudinal profiling due to BMI (n=36), diagnosis of diabetes (n=16), miscarriage (n=1) or withdrawal from the study (n=1). The final study population included 111 women (NBS n=70, RES n=16, MAL n=25) who contributed 491 maternal urine, 394 maternal blood, 68 maternal stool (online supplementary figure S1), 54 cord blood and 28 neonatal urine samples (online supplementary table S1). Two women (RES n=1, MAL n=1) contributed only neonatal samples. All women in the MAL group had a previous gastric bypass while the RES group had gastric banding (n=8) or sleeve gastrectomy (n=8). Maternal demographics, pregnancy characteristics and outcomes of the study population (online supplementary table S2) were consistent with our earlier report.[15] Compared with the NBS group, women with a previous BS were on average 3.7 years older and, especially those with a MAL procedure, had lower insulin resistance at T3 and delivered smaller babies earlier (online supplementary table S2). Among the BS group, women with a previous MAL procedure had a higher pre-surgery BMI compared with those with a previous RES procedure.

Metabolic differences in maternal and neonatal serum metabolomes associated with bariatric surgery

Clear differences between the serum metabolic profiles of MAL, but not RES, participants and NBS participants were detected in the third trimester of pregnancy (time points T4 and T5; online supplementary table S3) as evidenced by the OPLS-DA models. The serum metabolic profiles associated with BS at the time of labour and in the first and second trimesters could not be differentiated from the control group. Serum samples obtained from MAL patients contained lower concentrations of unsaturated lipids (CH=CH, δ 5.31 (m)) at T5 and the lipids group corresponding to $C{=}CCH_2C{=}C$ (δ 2.79 (m)) at both T4 and T5 (online supplementary figure S2) compared with the NBS group. Other metabolic differences between the MAL and NBS groups included (online supplementary table S4) *N*-acetyl glycoprotein (GlycA), leucine, isoleucine, isobutyrate (decreased in MAL) and glutamine, D-β-hydroxybutyrate (increased in MAL). The behaviour of these metabolites throughout the pregnancy period was investigated further using time series analysis to model individual metabolites. Serum leucine, isoleucine and isobutyrate concentrations were significantly lower in MAL (n=12; NBS n=34) throughout the time course (figure 1A), whereas increases in D-β-hydroxybutyrate were associated with the last few weeks of pregnancy and delivery. GlycA and glutamine were

Gut: first published as 10.1136/gutjnl-2019-319620 on 21 January 2020. Downloaded from http://gut.bmj.com/ on April 27, 2024 by guest. Protected by copyright.

not significantly changed in MAL during pregnancy. Significance of the differential serum metabolite concentrations at each time point is shown in online supplementary figure S3A. There was no significant difference between NBS and RES participants at any time point (online supplementary table S5). Although OPLS-DA and time series modelling showed clear evidence that previous BS impacted on the metabolic profiles, PCA indicated that the greatest metabolic variation in plasma composition was due to the changes over time during pregnancy, irrespective of surgery, with samples collected at T1 being markedly different from those obtained at later time points in the pregnancy (online supplementary figure S4). Thus no inherent structure associated with study group, age or BMI was evident in the ^{1}H NMR maternal serum data. No differences in the metabolic profiles of cord serum from neonates in each study group were detected (online supplementary table S3).

Surgery-associated metabolic differences in maternal and neonatal urine metabolomes

Similar to the serum samples, PCA models indicated that the metabolic variation driven by temporal changes throughout pregnancy was stronger than that associated with study group, age or BMI (online supplementary figure S4). Urine samples collected at delivery (T6) clearly differed from those taken at any other time point during pregnancy. In addition to endogenous changes, drug metabolites of paracetamol as well as mannitol, a common ingredient in drug tablet coating, were detected in urine obtained at T6. Outliers identified from PCA models at each time point were samples containing high concentrations of glucose or drug metabolites and were excluded from subsequent supervised models (online supplementary table S3). Based on the OPLS-DA models, no significant differences were identified between NBS and RES groups or between RES and MAL groups at any of the six time points; however, urine metabolic profiles of MAL patients were significantly altered compared with the NBS group as early as T1 and this difference persisted throughout pregnancy (online supplementary table S3). The strongest discriminatory metabolites (online supplementary table S4 and online supplementary figure S5) contributing to the differences between MAL and NBS groups were host-gut bacterial co-metabolites PAG, PCS, IS, PHPA, as well as an unidentified aromatic metabolite (unknown) with structural similarities to PCS (based on both ^{1}H and ^{13}C shifts; online supplementary figure S6). These metabolites were significantly elevated in MAL (n=14; NBS n=50) patients during pregnancy (figure 1A). Creatinine and α-ketoisovalerate were present in lower concentrations in MAL patients; methylmalonate and valine were not significantly associated with the MAL group over the time course (figure 1A) but were significantly lower in MAL patients at specific time points (online supplementary figure S3B). None of these metabolites were significantly changed in the RES group after correcting for multiple hypothesis tests (online supplementary table S5). Interestingly, when the RES group was divided into gastric banding (purely mechanical) and sleeve gastrectomy (removal of part of the stomach with metabolic effects), patients with a sleeve tended to have higher PAG and PCS than patients with a band (online supplementary figure S7). Maternal urine samples collected at T6 could not be discriminated for any pairwise comparison by OPLS-DA (online supplementary table S3). However PAG, the metabolite with the largest effect size in maternal samples, retained significance in the offspring's urine at T6 (figure 1B). The relative concentration of PAG was significantly increased in babies born from MAL mothers (n=4) compared with babies born from RES (n=6, p=0.04) and NBS (n=18, p=0.021) mothers, mirroring the difference observed between the MAL and NBS maternal samples.

Taxonomic changes in the maternal gut microbiota of MAL patients

To support the finding of altered microbiota-associated metabolites in MAL patients detected by ^{1}H NMR spectroscopy, faecal microbiome compositions from a subset of the cohort were analysed. A total of 68 stool samples representing 39 mothers (NBS n=25, MAL n=14) were studied. After excluding one maternal NBS sample due to low sequencing depth (2181 reads), the mean number of high-quality, paired-end 16S rRNA gene amplicon sequences per sample was 22 405 (±5535 SD) with a minimum sequencing depth of 12 967.

Women with a previous MAL surgery exhibited greater α-diversity, or a higher number of distinct taxa with more even distributions, compared with the NBS group but this was only statistically significant in the third trimester of pregnancy (figure 2A). β-diversity analysis comparing pairwise dissimilarities between samples revealed that MAL patients had a distinct (p=0.001) gut microbiota from the NBS group throughout pregnancy (figure 2B). MAL patients could be distinguished by an increase in relative abundance of *Escherichia/Shigella*, *Streptococcus* and *Enterococcus* genera and these changes manifested at all higher taxonomic ranks (figure 2C). Relative increases in abundances of *Rothia* and its family *Micrococcaceae* were also observed. *Anaerostipes* was the only genus found in significantly lower relative abundance in the MAL group at multiple time points. Only taxa that exhibited differential abundance in the MAL group at more than one time point were included in figure 2C, but taxa displaying differences between groups at a single time point during pregnancy are detailed in online supplementary table S6. Differential abundance at the amplicon sequence level was investigated for sequences assigned to the genera *Streptococcus*, *Enterococcus*, *Escherichia/Shigella* and *Rothia*. Species assignments were compared against SILVA[26] and RDB[27] databases. From the 25 16S rRNA gene sequences (*Escherichia/Shigella* n=11, *Streptococcus* n=12, *Enterococcus* n=1 and *Rothia* n=1) found in significantly higher abundance in MAL patients (online supplementary table S7), the following species were identified: *Escherichia coli*, *Streptococcus salivarius*, *Streptococcus vestibularis*, *Streptococcus mutans* and *Streptococcus parasanguinis*.

Integrative analysis of phenotypes

To determine if the observed MAL signatures were complementary across the data sets, an integrative approach was used to model the bacterial (n=68) and metabolic (serum n=60; urine n=66) data together. Clinical classes could be discriminated on the first component of the model for both urine and faecal data sets while this difference was not apparent in the serum data (figure 3A). Urine PAG, PCS, and IS concentrations were highly correlated and had the highest contribution to the model followed by leucine, isoleucine (also highly correlated) and isobutyrate (figure 3B). PAG, PCS and IS were most correlated with *Streptococcus*, *Enterococcus*, *Escherichia/Shigella*, *Rothia* and *Holdemanella* (figure 3C). Variables in this cluster also had a strong negative correlation with serum leucine, isoleucine and isobutyrate. *Anaerostipes*, the only genus that decreased in abundance in the MAL group, was negatively correlated with PAG, PCS and IS but positively correlated with leucine, isoleucine and isobutyrate (figure 3C).

Gut: first published as 10.1136/gutjnl-2019-319620 on 21 January 2020. Downloaded from http://gut.bmj.com/ on April 27, 2024 by guest. Protected by copyright.

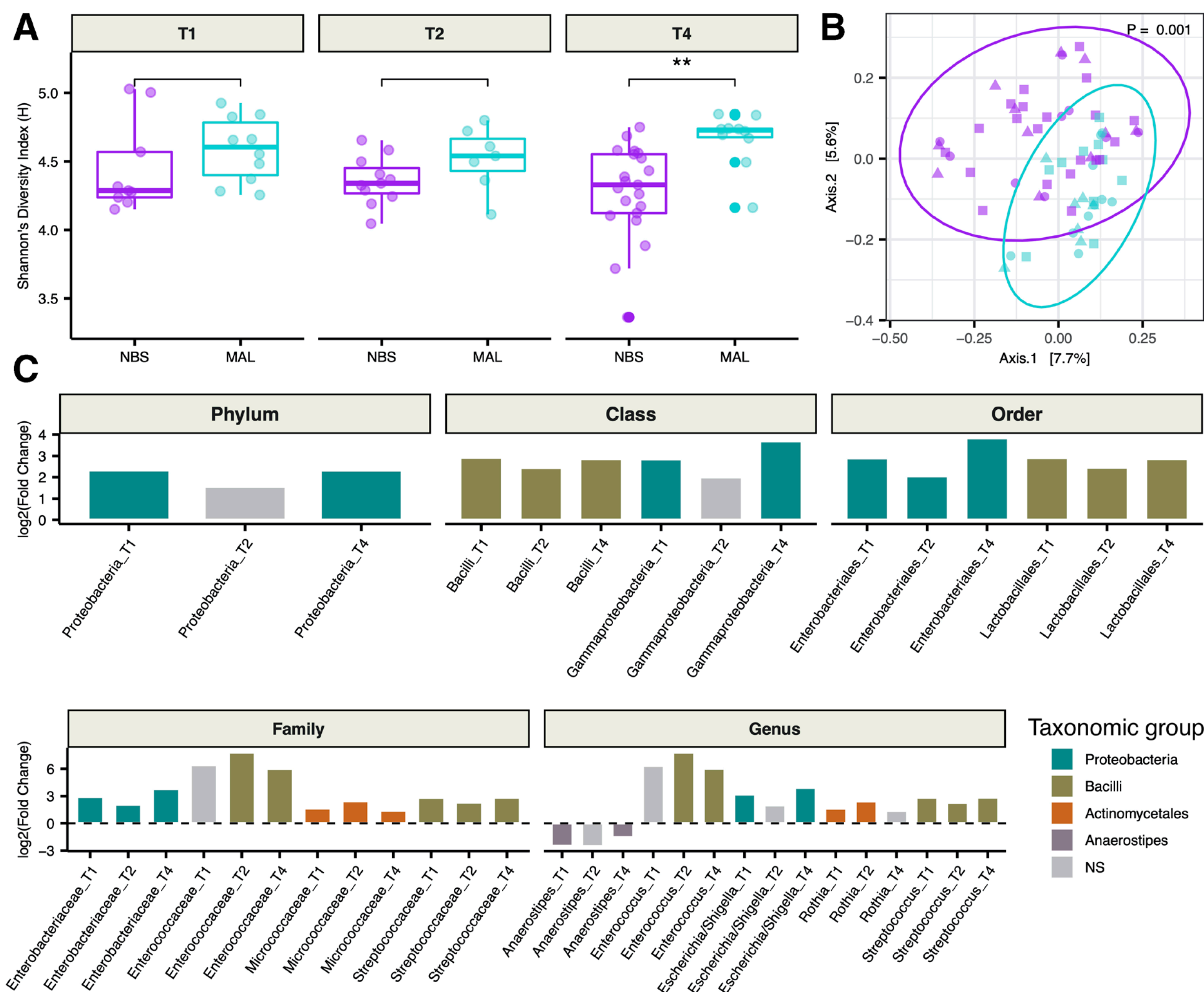

Figure 2 Gut bacterial alterations in malabsorptive patients. (A) Comparison of Shannon diversity between malabsorptive (MAL) and no bariatric surgery (NBS) groups at 11–14 weeks (T1), 20–24 weeks (T2) and 30–33 weeks (T4) gestation. (B) Principal coordinates analysis of Bray-Curtis dissimilarities between MAL (blue) and NBS (purple) during pregnancy (T1: circle, T2: triangle, T4: square). P value obtained from permutational multivariate analysis of variance test. (C) Taxa that exhibited differential abundance (Mann-Whitney U test; p_{adj}<0.05) at more than one time point are plotted. Values are reported in online supplementary table S6. **P<0.01 (Mann-Whitney U test); NS, not significant.

Metabolic profile and clinical outcomes

Spearman's correlation coefficients (ρ) were calculated to assess the relationships between relative concentrations of maternal metabolites and clinical measurements (maternal: fasting insulin, fasting glucose and HOMA-IR measured at T3; neonatal: estimated fetal weight (measured at T2, T4 and T5) or birth weight percentiles, and gestational age at delivery). Throughout pregnancy, urinary host-microbial co-metabolites PAG and PCS were negatively associated with maternal HOMA-IR and fasting insulin while the branched-chain amino acids (mainly at T2) had a positive correlation (figure 4). PAG and PCS relative concentrations at T2, IS and unknown at T5, and IS and PAG at T6 were also negatively correlated with the weight percentiles of the babies at the corresponding time points. Partial correlations controlling for the effect of maternal age and BMI on maternal measurements as well as the effect of maternal age, BMI and HOMA-IR on birthweight percentile are reported in online supplementary table S8, but these confounders did not alter the overall results. Importantly, the concentrations of these metabolites did not correlate (online supplementary figure S8) with percentage of weight lost or time interval between BS and conception (which ranged from 12 months to 11 years in our cohort) in MAL patients, suggesting that the phenotype is persistent.

DISCUSSION

Our results demonstrate that pregnancies following a MAL bariatric procedure are characterised by altered maternal and neonatal metabolic profiles compared with pregnancies without such surgery. In particular, women with previous MAL surgery had lower serum concentrations of branched-chain amino acids (leucine and isoleucine) and branched-chain fatty acids (isobutyrate) and excreted higher concentrations of urinary host-microbial co-metabolites of protein putrefaction (PAG, PCS, IS and PHPA). This signature was correlated with a shift in the gut microbiota which included increases in relative abundance of bacterial genera *Enterococcus*, *Streptococcus*, *Escherichia/Shigella* and *Rothia* as well as a decrease in *Anaerostipes*. The observed MAL-associated changes in metabolism were also inversely associated with maternal insulin resistance and the offspring's weight, suggesting that a previous MAL surgery may have both beneficial and detrimental effects on pregnancy. With regard to the RES group, we posit these patients, especially those with a previous sleeve gastrectomy, represent an intermediate phenotype due to the lower risk of malabsorption imposed by these procedures and less extreme changes to the gastrointestinal environment which would likely result in lesser alterations in the gut microbiome.[28]

Gut: first published as 10.1136/gutjnl-2019-319620 on 21 January 2020. Downloaded from http://gut.bmj.com/ on April 27, 2024 by guest. Protected by copyright.

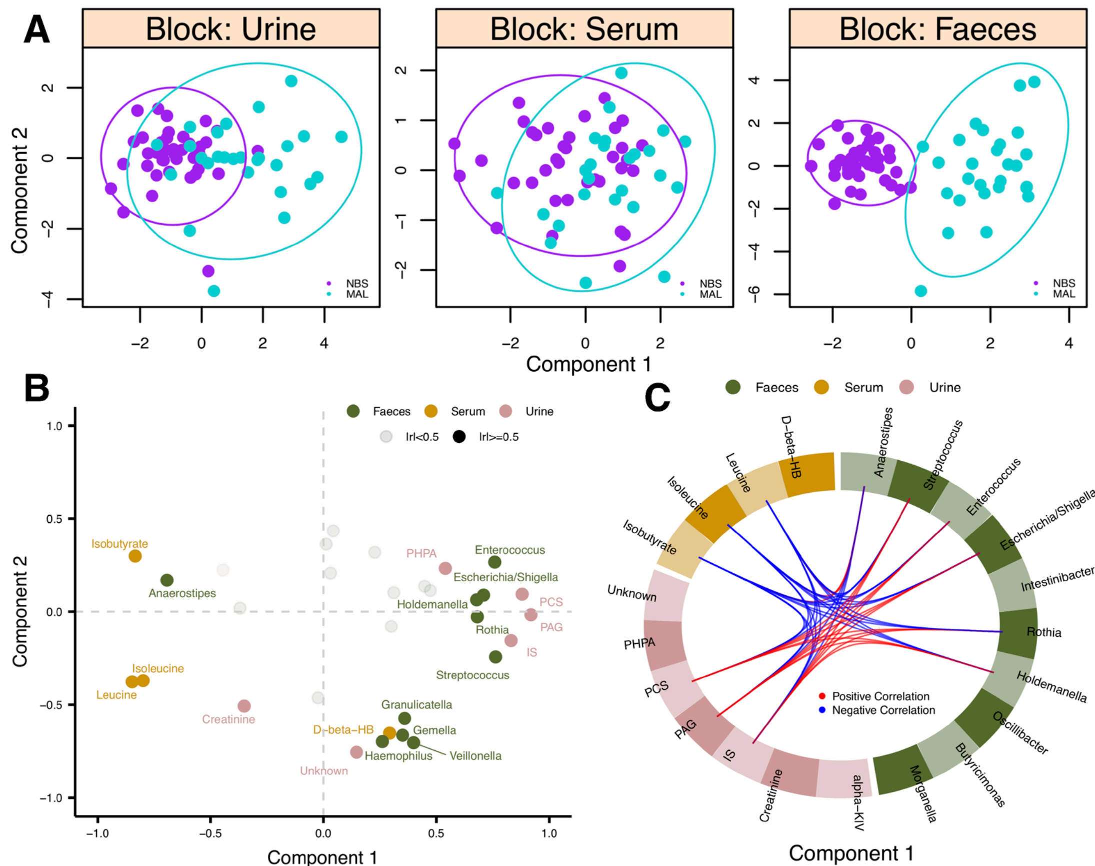

Figure 3 Integration of serum, urine and faecal profiles. (A) Scores plots showing samples from each data set projected in latent space. Clinical classes are discriminated along component 1. (B) Variables are plotted (coloured by biofluid) where each point represents the Pearson correlation between that variable and each component. High correlations with either component (|r|≥0.5) are BOLD and labelled. (C) Pearson correlations (|r|≥0.5) between variables in different data sets along component 1 are plotted. NBS, no bariatric surgery; MAL, malabsorptive; alpha-KIV, α-ketoisovalerate; D-beta-HB, D-β-hydroxybutyrate; IS, indoxyl sulfate; PAG, phenylacetylglutamine; PCS, *p*-cresol sulfate; PHPA, *p*-hydroxyphenylacetate.

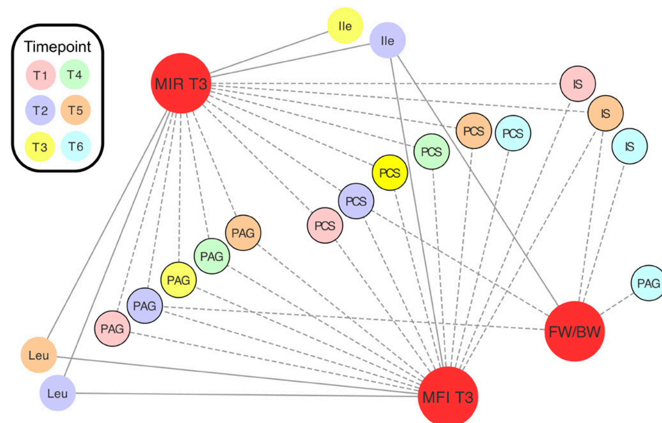

Figure 4 Relative concentrations of metabolites correlate with maternal insulin resistance and fetal/birth weight. Spearman's correlations between clinical measurements (red) and metabolite relative concentrations. Correlations plotted are significant (p_{adj}<0.05). Values are reported in online supplementary table S8. Negative correlations are denoted by dashed lines. Urine metabolites are denoted by black borders. FW/BW, estimated fetal weight or birth weight; Ile, isoleucine; IS, indoxyl sulfate; Leu, leucine; MFI, maternal fasting insulin; MIR, maternal HOMA-IR; PAG, phenylacetylglutamine; PCS, *p*-cresol sulfate.

Our results assessing the influence of BS on the maternal faecal microbiome composition are consistent with previous findings in non-pregnant individuals where MAL, but not RES, surgery resulted in increased relative abundances of *Bacilli* and *Gammaproteobacteria* (including genera *Enterococcus*, *Streptococcus* and *Escherichia*) as well as associated faecal metabolites[28]; abundances of both classes positively correlated with products of protein fermentation. Studies consistently show a gut bacterial signature of facultative anaerobes and oral-associated bacteria in BS (mainly MAL) patients with an increase in relative abundance of enterobacteria being the most consistent bacterial marker across multiple studies in both humans[7 8 10 28–30] and animals.[11 31] This shift in gut microbial community structure is believed to be, at least partially, a result of the increased oxygen and pH in the digestive tract following the anatomical rearrangement of a MAL surgery.[30 32] Enterobacteria are Gram-negative and generally present in very low densities (<<10^8 CFU/g) in the normal gut, but increased abundances of these bacteria, some of which are known pathogens, have been associated with both inflammatory bowel disease and colon cancer.[33] This underscores the need for understanding the mechanisms by which the altered bariatric microbiome can impact on maternal and neonatal metabolism and downstream health.

Increased concentrations of urinary microbial-associated metabolites following MAL BS have been reported in animals[11 34] and in one human study with only two bariatric patients.[9] Carbohydrates are the preferred source of energy for

Gut: first published as 10.1136/gutjnl-2019-319620 on 21 January 2020. Downloaded from http://gut.bmj.com/ on April 27, 2024 by guest. Protected by copyright.

microbes inhabiting the large intestine.[35] However, we hypothesise that the MAL condition not only changes the gastrointestinal environment resulting in an altered microbial community but also results in more undigested protein reaching the large intestine[36] and both of these factors likely contribute to the shift from carbohydrate to protein metabolism. Fermentation of protein by gut bacteria, or putrefaction, in the colon is generally considered detrimental to health.[37] Gut bacterial metabolism of aromatic amino acids (phenylalanine, tyrosine and tryptophan) results in the production of phenols and indoles,[38] which are inflammatory and potentially toxic compounds.[39 40] Bacterial enzymes are necessary for metabolism of phenylalanine and tyrosine to phenolic compounds phenylacetate and *p*-cresol, respectively, and for the metabolism of tryptophan to indole. Indole and *p*-cresol are sulfated to IS and PCS, respectively, prior to excretion while phenylacetate is converted to PAG by host metabolism. PHPA is an intermediate in the tyrosine degradation pathway but is also formed from phenylacetate.[38] *Enterococcus* spp. contain enzymes necessary for both phenylalanine and tyrosine metabolism; *E. coli* is involved in tyrosine metabolism.[41] Both species are also known as indole producers.[42] Although *Streptococcus* may not be directly involved in the formation of these metabolites, it does have proteolytic activity[43] and could contribute to the observed phenotype by making amino acids available from dietary protein. To our knowledge, this is the first report of upregulation of host-microbe co-metabolism of aromatic amino acids post-MAL surgery in humans. Previous reports have shown that the gut microbiota shifts to protein putrefaction after a MAL BS[10 28 29] but, having focused on faecal samples, the host component of these pathways was overlooked. Lacking dietary data, we cannot exclude the possibility that the MAL metabolic phenotype results from increased protein intake. However, reports of protein malnutrition and protein intolerance post-BS[44 45] suggest that high protein intake is unlikely in these patients. An inverse association between PAG/PCS and BMI has been reported although the relationship between adipiosity and the gut microbiota is still under debate.[46] Other reports, in agreement with our data, have identified decreased concentrations of branched-chain amino acids in non-pregnant BS patients which supports our hypothesis that protein metabolism is altered in these individuals. Branched-chain amino acids have previously been linked to insulin resistance associated with obesity.[47–50]

Our data suggest that PAG and PCS, or their respective metabolic pathways, are also associated with reduction in maternal insulin resistance following BS. We are aware of recent discussion surrounding the suitability of oral glucose tolerance testing in bariatric patients and have since revised our methods. In spite of this caveat, it is conceivable that the gut microbiota plays a role in insulin resistance.[49 51] Our finding of an inverse correlation between maternal urinary host-microbial co-metabolites and fetal/birth weight indicates that the maternal metabolic changes may also be related to the risk of these women delivering small for gestational age neonates. Although malnutrition resulting from maternal malabsorption likely contributes to reduced fetal growth, the downstream metabolic consequences of malabsorption also warrant further investigation given the potential toxicity of the final metabolic products. It has also been shown that higher urinary concentrations of PCS are indicative of reduced sulfation capacity in the host, which presents significant competition for similar reactions, namely acetaminophen detoxification, which also require sulfation prior to excretion.[52] Pharmacokinetics in these patients should be investigated, especially during pregnancy, given that MAL patients have higher urinary concentrations of two sulfated compounds (PCS and IS).

Pregnancy has been shown to induce substantial time-dependent alterations in physiology and metabolism[53 54] under the constantly changing physiological demand as the mother responds to the needs of the growing fetus. Nevertheless, the systematic effect of BS on the metabolic profiles is superimposed on the changing metabolic landscape and can be clearly observed in the metabolic phenotypes. In particular, GlycA, a metabolite associated with systemic inflammation, increased during the first two trimesters but remained lower in MAL patients during the third trimester indicating that these women have a lower grade of pregnancy-related inflammation compared with obese women without BS.[50 55] Drastic metabolic changes could also be seen at the time of delivery in all groups which is likely due to the metabolic consequences of the labour process and masked surgery-related differences.[56] Intriguingly, despite prophylactic administration of antibiotics (given to 59% of the women at delivery), PAG was detected in the neonatal urine soon after birth and in significantly higher concentration in those born from women with previous MAL surgery compared with the offspring of the NBS women. Although this result should be treated with caution due to the low number of newborns enrolled in the study, it suggests transfer of the modified maternal profile to the offspring. Further independent studies are required to validate the generalisability of this observation with longer-term follow-up of the infants to determine whether this neonatal metabolic phenotype is maintained and how it impacts their future risk of obesity and diabetes. BS is not an appropriate intervention at population scale but identification of beneficially altered physiology may allow for development of targeted interventions aimed at specifically modulating key pathways.

Twitter Kiana Ashley West @kiana_a_west

Acknowledgements The authors would like to thank Dr Ivan Serrano Contreras for assistance with carrying out 2D NMR spectroscopy. We would also like to acknowledge that KAW is supported by the STRATiGRAD doctoral training program and JVL is supported by Medical Research Council New Investigator Grant (P62104) and European Research Council Starting Grant (P61481). EH is supported by the Western Australian Government through the Premier's Science Fellowship Program. Infrastructure support was provided by the National Institute for Health Research Imperial Biomedical Research Centre aligned with the gut health theme.

Contributors EH, MDS, JVL and MRJ designed the study. CK and TM recruited patients and collected biospecimens. KAW, CK and JAKM performed the experiments. KAW analysed the data. All authors contributed to the interpretation of results. KAW, MDS and EH wrote the manuscript. All authors read and approved the final manuscript.

Funding The study was supported partially by BORNE (UK Registered Charity number: 1167073).

Disclaimer The funding body did not contribute to the design of the study, collection, analysis, interpretation of data or writing the manuscript.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study was approved by the West London Research Ethics Committee (No: 14/LO/0592).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Raw sequencing data are available in a public, open access repository: Sequence Read Archive under BioProject PRJNA493625. Raw NMR data are available upon reasonable request. All pre-processed data and code to reproduce the results are available in a public, open access repository: GitHub (https://github.com/ka-west/PBS_manuscript). The R script is also included as a pdf file created with 'knitr' (see online supplementary methods).

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is

Gut: first published as 10.1136/gutjnl-2019-319620 on 21 January 2020. Downloaded from http://gut.bmj.com/ on April 27, 2024 by guest. Protected by copyright.

properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Kiana Ashley West http://orcid.org/0000-0002-1896-7936
Julie Anne Kathryn McDonald http://orcid.org/0000-0003-0739-6047

REFERENCES

1 World Health Organization. *Obesity and overweight*, 2017.
2 Abdelaal M, le Roux CW, Docherty NG. Morbidity and mortality associated with obesity. *Ann Transl Med* 2017;5.
3 Ashrafian H, Ahmed K, Rowland SP, *et al*. Metabolic surgery and cancer. *Cancer* 2011;117:1788–99.
4 Noria SF, Grantcharov T. Biological effects of bariatric surgery on obesity-related comorbidities. *Can J Surg* 2013;56:47–57.
5 Ashrafian H, Athanasiou T, Li JV, *et al*. Diabetes resolution and hyperinsulinaemia after metabolic Roux-en-Y gastric bypass. *Obes Rev* 2011;12:e257–72.
6 Elder KA, Wolfe BM. Bariatric surgery: a review of procedures and outcomes. *Gastroenterology* 2007;132:2253–71.
7 Graessler J, Qin Y, Zhong H, *et al*. Metagenomic sequencing of the human gut microbiome before and after bariatric surgery in obese patients with type 2 diabetes: correlation with inflammatory and metabolic parameters. *Pharmacogenomics J* 2013;13:514–22.
8 Kong L-C, Tap J, Aron-Wisnewsky J, *et al*. Gut microbiota after gastric bypass in human obesity: increased richness and associations of bacterial genera with adipose tissue genes. *Am J Clin Nutr* 2013;98:16–24.
9 Calvani R, Miccheli A, Capuani G, *et al*. Gut microbiome-derived metabolites characterize a peculiar obese urinary metabotype. *Int J Obes* 2010;34:1095–8.
10 Tremaroli V, Karlsson F, Werling M, *et al*. Roux-En-Y gastric bypass and vertical banded gastroplasty induce long-term changes on the human gut microbiome contributing to fat mass regulation. *Cell Metab* 2015;22:228–38.
11 Li JV, Ashrafian H, Bueter M, *et al*. Metabolic surgery profoundly influences gut microbial-host metabolic cross-talk. *Gut* 2011;60:1214–23.
12 Nicholson JK, Holmes E, Kinross J, *et al*. Host-gut microbiota metabolic interactions. *Science* 2012;336:1262–7.
13 Holmes E, Li JV, Athanasiou T, *et al*. Understanding the role of gut microbiome-host metabolic signal disruption in health and disease. *Trends Microbiol* 2011;19:349–59.
14 Kwong W, Tomlinson G, Feig DS. Maternal and neonatal outcomes after bariatric surgery; a systematic review and meta-analysis: do the benefits outweigh the risks? *Am J Obstet Gynecol* 2018;218:573–80.
15 Maric T, Kanu C, Johnson MR, *et al*. Maternal, neonatal insulin resistance and neonatal anthropometrics in pregnancies following bariatric surgery. *Metabolism* 2019;97:25–31.
16 Dona AC, Jiménez B, Schäfer H, *et al*. Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Anal Chem* 2014;86:9887–94.
17 Cloarec O, Dumas M-E, Craig A, *et al*. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Anal Chem* 2005;77:1282–9.
18 Posma JM, Garcia-Perez I, De Iorio M, *et al*. Subset optimization by reference matching (storm): an optimized statistical approach for recovery of metabolic biomarker structural information from 1H NMR spectra of biofluids. *Anal Chem* 2012;84:10694–701.
19 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2017. https://www.R-project.org/
20 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc* 1995;57B:289–300.
21 Mullish BH, Pechlivanis A, Barker GF, *et al*. Functional microbiomics: evaluation of gut microbiota-bile acid metabolism interactions in health and disease. *Methods* 2018;149:49–58.
22 Callahan BJ, Sankaran K, Fukuyama JA, *et al*. Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *Version 2. F1000Res* 2016;5:1492.
23 Callahan BJ, McMurdie PJ, Rosen MJ, *et al*. DADA2: high-resolution sample inference from illumina amplicon data. *Nat Methods* 2016;13:581–3.
24 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol* 2014;15:550.
25 Rohart F, Gautier B, Singh A, *et al*. mixOmics: An R package for ʼomics feature selection and multiple data integration. *PLoS Comput Biol* 2017;13:e1005752.
26 Quast C, Pruesse E, Yilmaz P, *et al*. The Silva ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;41:D590–6.
27 Cole JR, Wang Q, Fish JA, *et al*. Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 2014;42:D633–42.
28 Ilhan ZE, DiBaise JK, Isern NG, *et al*. Distinctive microbiomes and metabolites linked with weight loss after gastric bypass, but not gastric banding. *ISME J* 2017;11:2047–58.
29 Palleja A, Kashani A, Allin KH, *et al*. Roux-En-Y gastric bypass surgery of morbidly obese patients induces swift and persistent changes of the individual gut microbiota. *Genome Med* 2016;8:67.
30 Zhang H, DiBaise JK, Zuccolo A, *et al*. Human gut microbiota in obesity and after gastric bypass. *Proc Natl Acad Sci U S A* 2009;106:2365–70.
31 Liou AP, Paziuk M, Luevano J-M, *et al*. Conserved shifts in the gut microbiota due to gastric bypass reduce host weight and adiposity. *Sci Transl Med* 2013;5:178ra41.
32 Aron-Wisnewsky J, Doré J, Clement K. The importance of the gut microbiota after bariatric surgery. *Nat Rev Gastroenterol Hepatol* 2012;9:590–8.
33 Liu Z, Cao AT, Cong Y. Microbiota regulation of inflammatory bowel disease and colorectal cancer. *Semin Cancer Biol* 2013;23:543–52.
34 Seyfried F, Li JV, Miras AD, *et al*. Urinary phenotyping indicates weight loss-independent metabolic effects of Roux-en-Y gastric bypass in mice. *J Proteome Res* 2013;12:1245–53.
35 Conlon M, Bird A. The impact of diet and lifestyle on gut microbiota and human health. *Nutrients* 2015;7:17–44.
36 Ponsky TA, Brody F, Pucci E. Alterations in gastrointestinal physiology after Roux-en-Y gastric bypass. *J Am Coll Surg* 2005;201:125–31.
37 Windey K, De Preter V, Verbeke K. Relevance of protein fermentation to gut health. *Mol Nutr Food Res* 2012;56:184–96.
38 Smith EA, Macfarlane GT. Formation of phenolic and indolic compounds by anaerobic bacteria in the human large intestine. *Microb Ecol* 1997;33:180–8.
39 Meyer TW, Hostetter TH. Uremic solutes from colon microbes. *Kidney Int* 2012;81:949–54.
40 Sun C-Y, Hsu H-H, Wu M-S. P-Cresol sulfate and indoxyl sulfate induce similar cellular inflammatory gene expressions in cultured proximal renal tubular cells. *Nephrol Dial Transplant* 2013;28:70–8.
41 Gryp T, Vanholder R, Vaneechoutte M, *et al*. p-Cresyl sulfate. *Toxins (Basel)* 2017;9:52.
42 Lee J-H, Lee J. Indole as an intercellular signal in microbial communities. *FEMS Microbiol Rev* 2010;34:426–44.
43 Macfarlane GT, Cummings JH, Allison C. Protein degradation by human intestinal bacteria. *Microbiology* 1986;132:1647–56.
44 Steenackers N, Gesquiere I, Matthys C. The relevance of dietary protein after bariatric surgery: what do we know? *Curr Opin Clin Nutr Metab Care* 2018;21:58–63.
45 Lupoli R, Lembo E, Saldalamacchia G, *et al*. Bariatric surgery and long-term nutritional issues. *World J Diabetes* 2017;8:464–74.
46 Elliott P, Posma JM, Chan Q, *et al*. Urinary metabolic signatures of human adiposity. *Sci Transl Med* 2015;7:285ra62-ra62.
47 Newgard CB, An J, Bain JR, *et al*. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab* 2009;9:311–26.
48 Tai ES, Tan MLS, Stevens RD, *et al*. Insulin resistance is associated with a metabolic profile of altered protein metabolism in Chinese and Asian-Indian men. *Diabetologia* 2010;53:757–67.
49 Pedersen HK, Gudmundsdottir V, Nielsen HB, *et al*. Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* 2016;535:376–81.
50 Houttu N, Mokkala K, Laitinen K. Overweight and obesity status in pregnant women are related to intestinal microbiota and serum metabolic and inflammatory profiles. *Clin Nutr* 2018;37:1955–66.
51 Kootte RS, Levin E, Salojärvi J, *et al*. Improvement of insulin sensitivity after lean donor feces in metabolic syndrome is driven by baseline intestinal microbiota composition. *Cell Metab* 2017;26:611–9.
52 Clayton TA, Baker D, Lindon JC, *et al*. Pharmacometabonomic identification of a significant Host-microbiome metabolic interaction affecting human drug metabolism. *Proc Natl Acad Sci U S A* 2009;106:14728–33.
53 Liu X, Wang X, Sun H, *et al*. Urinary metabolic variation analysis during pregnancy and application in gestational diabetes mellitus and spontaneous abortion biomarker discovery. *Sci Rep* 2019;9:2605.
54 Wang M, Xia W, Li H, *et al*. Normal pregnancy induced glucose metabolic stress in a longitudinal cohort of healthy women: novel insights generated from a urine metabolomics study. *Medicine* 2018;97:e12417.
55 Otvos JD, Shalaurova I, Wolak-Dinsmore J, *et al*. GlycA: a composite nuclear magnetic resonance biomarker of systemic inflammation. *Clin Chem* 2015;61:714–23.
56 Caboni P, Meloni A, Lussu M, *et al*. Urinary metabolomics of pregnant women at term: a combined GC/MS and NMR approach. *J Matern Fetal Neonatal Med* 2014;27(Suppl 2):4–12.

Gut: first published as 10.1136/gutjnl-2019-319620 on 21 January 2020. Downloaded from http://gut.bmj.com/ on April 27, 2024 by guest. Protected by copyright.

SUPPLEMENTARY DATA

Longitudinal metabolic and gut bacterial profiling of pregnant women with previous bariatric surgery

Contents Page

Figure S1: Summary of sampling by individual..................... 2
Table S1: Samples analysed..................... 2
Table S2: Clinical characteristics and pregnancy outcomes of the study participants..................... 3
Table S3: Summary of OPLS-DA model statistics..................... 4
Figure S2: Overlayed median 1H NMR serum spectra..................... 5
Table S4: Metabolite identification..................... 6
Figure S3: Maternal metabolic changes depend on type of bariatric surgery..................... 7
Table S5: Maternal metabolic changes depend on type of bariatric surgery..................... 8
Figure S4: Principal component analysis (PCA) of 1H NMR spectral data..................... 10
Figure S5: Overlayed median 1H NMR urine spectra..................... 11
Figure S6: Chemical shifts of unknown metabolite..................... 11
Figure S7: Relative concentrations of urinary host-microbial co-metabolites vary depending on bariatric surgery sub-type..................... 12
Table S6: Differentially abundant taxa in malabsorptive group..................... 13
Table S7: Differentially abundant amplicon sequence variants..................... 15
Table S8: Relative concentrations of metabolites correlate with maternal insulin resistance and fetal/birth weight..................... 16
Figure S8: Key metabolites are not associated with percentage of weight lost or length of time since surgery in malabsorptive patients..................... 17

Figure S1: Summary of sampling by individual. Each point represents a sampling event for an individual and is coloured based on which samples were collected and analysed. The size of the point represents the BMI category of the individual at that time point. *T1 11-14 weeks gestation; T2 20-24 weeks gestation; T3 28-30 weeks gestation; T4 30-33 weeks gestation; T5 35-37 weeks gestation; T6 delivery*

Table S1: Samples analysed. Number of samples analysed for each biospecimen type at each time point.

	Group	Maternal							Infant
		T1	T2	T3	T4	T5	T6	Total	T6
Serum	NBS	54	47	47	43	38	27	256	30
	RES	6	10	12	8	10	7	53	10
	MAL	12	15	15	15	17	11	85	14
Urine	NBS	63	61	58	52	51	41	326	18
	RES	6	11	11	12	12	12	64	6
	MAL	16	16	17	18	18	16	101	4
Faeces	NBS	9	11	-	21	-	-	41	-
	MAL	10	7	-	10	-	-	27	-

NBS no bariatric surgery; RES restrictive; MAL malabsorptive; T1 11-14 weeks gestation; T2 20-24 weeks gestation; T3 28-30 weeks gestation; T4 30-33 weeks gestation; T5 35-37 weeks gestation; T6 delivery

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

Table S2: Clinical characteristics and pregnancy outcomes of the study participants

Variable	No bariatric surgery (n = 70)	Post-bariatric surgery (n = 41)	Restrictive (n = 16)	Malabsorptive (n = 25)
Maternal age, years	29.72 ± 5.26	33.46 ± 4.58*	32.56 ± 4.32*	34.04 ± 4.8*
Parity, n (%)				
Nulliparous	43 (61.4)	20 (48.8)	9 (56.3)	11 (44.0)
Parous	27 (38.6)	21 (51.2)	7 (43.8)	14 (56.0)
Racial group, n (%)				
White	54 (77.1)	31 (75.6)	12 (75)	19 (76.0)
Other	16 (22.9)	10 (24.4)	4 (25)	6 (24.0)
Conception, n (%)				
Spontaneous	67 (95.7)	38 (92.7)	16 (100.0)	22 (88.0)
Assisted reproductive techniques	3 (4.3)	3 (7.3)	0 (0)	3 (12.0)
Smoking, n (%)				
No	67 (95.7)	36 (87.8)	15 (93.8)	21 (84.0)
Yes	3 (4.3)	5 (12.2)	1 (6.3)	4 (16.0)
Time between surgery and conception, months	-	58.34 ± 31.24	49.12 ± 30.54	64.24 ± 30.83ɫ
BMI prior to surgery, kg/m²	-	47.19 ± 7.95	41.95 ± 6.47	50.54 ± 7.01ɫ
Booking BMI, kg/m²	34.12 ± 5.68	32.94 ± 5.10	32.96 ± 7.29	32.93 ± 4.10
Gestational age at delivery, weeks	39.44 ± 1.34	38.67 ± 2.266*	38.95 ± 2.31	38.49 ± 2.26*
Mode of delivery, n (%)				
Vaginal	38 (54.3)	21 (51.2)	6 (37.5)	15 (60.0)
Caesarean section	32 (45.7)	20 (48.8)	10 (62.5)	10 (40.0)
Birth weight, g	3520.92 ± 603.86	3062.53 ± 588.32*	3096.75 ± 519.07*	3040.64 ± 638.12*
Birth weight percentile	59.48 ± 33.30	39.01 ± 27.35*	39.29 ± 26.25*	38.83 ± 28.57*

*Data are expressed as mean ± standard deviation or as otherwise stated. *P <0.05 comparison to no bariatric surgery group; ɫP <0.05 comparison between restrictive and malabsorptive groups*

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

Table S3: Summary of OPLS-DA model statistics. OPLS-DA model statistics for two-group comparisons are reported for each time point after excluding extreme outliers identified from PCA models.

Time point	Biofluid	Samples	Outliers Excluded	Class	R^2X	R^2Y	Q^2	*P* value
11-14 weeks (T1)	Urine	NBS n = 63	NBS n = 3	NBS vs RES	0.119	0.776	-0.117	NS
		RES n = 6	RES n = 1	RES vs MAL	0.171	0.935	-0.025	NS
		MAL n = 16	MAL n = 2	**NBS vs MAL**	**0.088**	**0.830**	**0.159**	**0.015388**
	Serum	NBS n = 54	NBS n = 0	NBS vs RES	0.022	0.817	-0.274	NS
		RES n = 6	RES n = 0	RES vs MAL	0.058	0.929	-0.044	NS
		MAL n = 12	MAL n = 0	NBS vs MAL	0.037	0.749	-0.156	NS
20-24 weeks (T2)	Urine	NBS n = 61	NBS n = 1	NBS vs RES	0.096	0.810	0.094	NS
		RES n = 11	RES n = 1	RES vs MAL	0.140	0.941	0.297	NS
		MAL n = 16	MAL n = 0	**NBS vs MAL**	**0.095**	**0.843**	**0.324**	**1.17E-05**
	Serum	NBS n = 47	NBS n = 0	NBS vs RES	0.027	0.830	-0.067	NS
		RES n = 10	RES n = 0	RES vs MAL	0.051	0.900	0.056	NS
		MAL n = 15	MAL n = 0	NBS vs MAL	0.031	0.825	0.114	NS
28-30 weeks (T3)	Urine	NBS n = 58	NBS n = 0	NBS vs RES	0.111	0.826	0.046	NS
		RES n = 11	RES n = 1	RES vs MAL	0.187	0.893	0.343	NS
		MAL n = 17	MAL n = 3	**NBS vs MAL**	**0.100**	**0.843**	**0.407**	**3.71E-07**
	Serum	NBS n = 47	NBS n = 0	NBS vs RES	0.033	0.748	-0.218	NS
		RES n = 12	RES n = 0	RES vs MAL	0.049	0.948	0.216	NS
		MAL n = 15	MAL n = 0	NBS vs MAL	0.040	0.744	0.091	NS
30-33 weeks (T4)	Urine	NBS n = 52	NBS n = 2	NBS vs RES	0.121	0.790	0.052	NS
		RES n = 12	RES n = 1	RES vs MAL	0.131	0.921	0.086	NS
		MAL n = 18	MAL n = 1	**NBS vs MAL**	**0.119**	**0.812**	**0.344**	**2.42E-05**
	Serum	NBS n = 43	NBS n = 0	NBS vs RES	0.034	0.742	-0.098	NS
		RES n = 8	RES n = 0	RES vs MAL	0.075	0.881	0.214	NS
		MAL n = 15	MAL n = 0	**NBS vs MAL**	**0.042**	**0.775**	**0.178**	**0.027**
35-37 weeks (T5)	Urine	NBS n = 51	NBS n = 2	NBS vs RES	0.117	0.802	0.023	NS
		RES n = 12	RES n = 0	RES vs MAL	0.135	0.935	0.001	NS
		MAL n = 18	MAL n = 0	**NBS vs MAL**	**0.109**	**0.852**	**0.483**	**2.11E-08**
	Serum	NBS n = 38	NBS n = 0	NBS vs RES	0.026	0.908	-0.377	NS
		RES n = 10	RES n = 0	RES vs MAL	0.054	0.916	0.224	NS
		MAL n = 17	MAL n = 0	**NBS vs MAL**	**0.048**	**0.847**	**0.245**	**0.006**
Delivery (T6)	Urine	NBS n = 41	NBS n = 2	NBS vs RES	0.087	0.844	0.055	NS
		RES n = 12	RES n = 0	RES vs MAL	0.164	0.907	0.110	NS
		MAL n = 16	MAL n = 4	NBS vs MAL	0.134	0.767	0.067	NS
	Serum	NBS n = 27	NBS n = 0	NBS vs RES	0.035	0.895	-0.116	NS
		RES n = 7	RES n = 0	RES vs MAL	0.068	0.903	-0.103	NS
		MAL n = 11	MAL n = 0	NBS vs MAL	0.055	0.789	0.028	NS
	Cord Blood	NBS n = 30	NBS n = 0	NBS vs RES	0.030	0.886	-0.022	NS
		RES n =10	RES n =0	RES vs MAL	0.043	0.925	-0.025	NS
		MAL n = 14	MAL n = 0	NBS vs MAL	0.037	0.885	0.031	NS
	Newborn urine	NBS n = 18	NBS n = 0	Univariate analysis only				
		RES n = 6	RES n = 0					
		MAL n = 4	MAL n = 0					

NBS no bariatric surgery; RES restrictive; MAL malabsorptive; NS not significant (P<0.05)

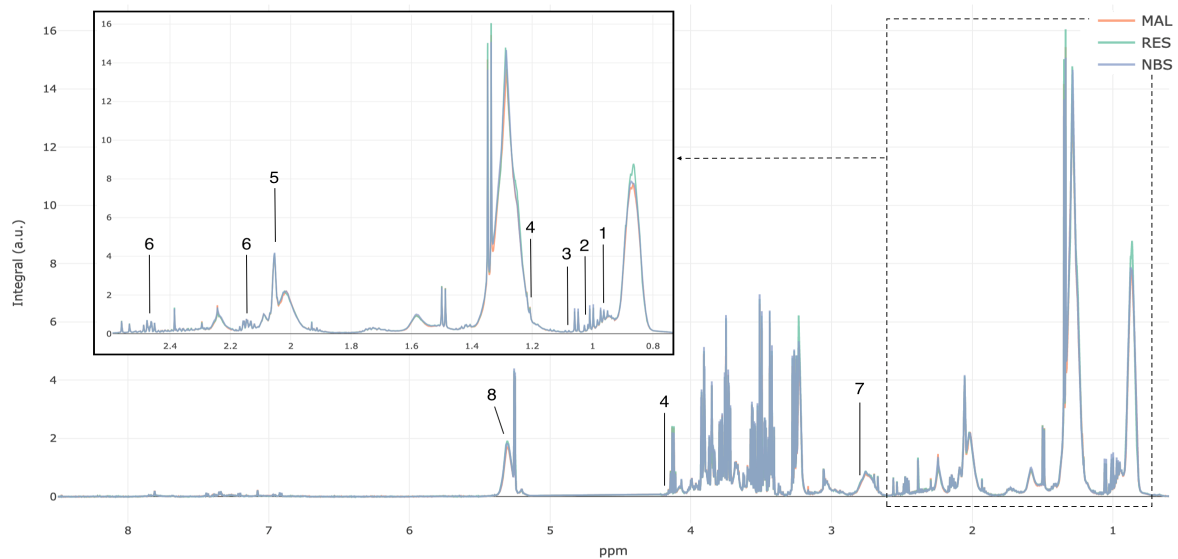

Figure S2: Overlayed median ^{1}H NMR serum spectra. A median spectrum is plotted for each study group. Annotated peaks and peak groups are those that were considered discriminatory in OPLS-DA models comparing malabsorptive (MAL) and no bariatric surgery (NBS) patients. The representative spectrum for restrictive (RES) patients is also plotted. 1 leucine; 2 isoleucine; 3 isobutyrate; 4 D-β-hydroxybutyrate ; 5 N-acetyl glycoprotein (GlycA) ; 6 glutamine ; 7 lipids (C=CCH$_2$C=C) ; 8 unsaturated lipids (CH=CH)

Table S4: Metabolite identification

Metabolite	^{1}H NMR Chemical Shifts*	Confirmed by	Biofluid
Leucine	0.96 (2xd), 1.71 (m), 3.73 (t)	STOCSY/STORM, J-res	Serum
Isoleucine	0.94 (t), 1.01 (d)	STOCSY/STORM, J-res	Serum
Isobutyrate	1.09 (d)	STOCSY/STORM, J-res, HSQC	Serum
N–acetyl glycoprotein	2.05 (s)	STOCSY/STORM, J-res, HSQC	Serum
Glutamine	2.14 (m), 2.46 (m), 3.77 (t)	STOCSY/STORM, J-res, HSQC	Serum
D-β-hydroxybutyrate	1.22 (d), 2.37 (m), 4.17 (m)	STOCSY/STORM, J-res	Serum
Valine	0.99 (d), 1.04 (d)	STOCSY/STORM, J-res	Urine
α-ketoisovalerate	1.11 (d)	STOCSY/STORM, J-res	Urine
Methylmalonate	1.24 (d)	STOCSY/STORM, J-res	Urine
Creatinine	3.05 (s), 4.06 (s)	STOCSY/STORM, J-res	Urine
p-cresol sulfate	2.35 (s), 7.21 (d), 7.28 (d)	STOCSY/STORM, J-res, HSQC	Urine
p-hydroxyphenylacetate	3.45 (s), 6.86 (d), 7.17 (d)	STOCSY/STORM, J-res, HSQC, spike in	Urine
Phenylacetylglutamine	1.92 (m), 2.11 (m), 2.27 (m), 3.67 (m), 4.19 (m), 7.36 (t), 7.43 (t)	STOCSY/STORM, J-res, HSQC	Urine
Unknown	7.34 (d), 7.48 (t)	STOCSY/STORM, J-res, HSQC	Urine
Indoxyl sulfate	7.5 (d), 7.7 (d)	STOCSY/STORM, J-res, HSQC, spike in	Urine

**s=singlet, d=doublet, t=triplet, m=multiplet*

STOCSY: Statistical Total Correlation Spectroscopy; STORM: Subset Optimisation by Reference Matching; HSQC: Hetero-nuclear Single Quantum Coherence (2D NMR)

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

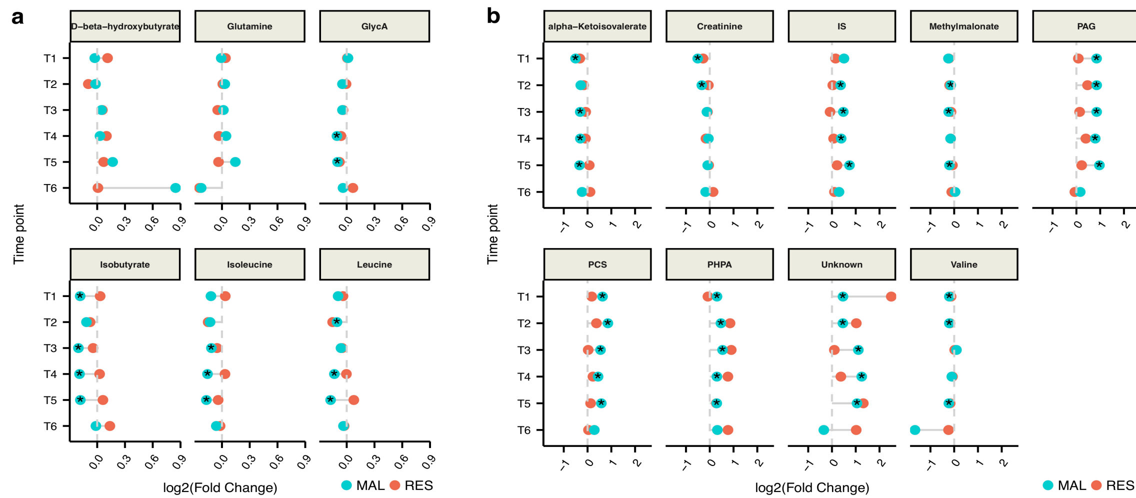

Figure S3: Maternal metabolic changes depend on type of bariatric surgery. For each discriminatory metabolite identified in (a) serum or (b) urine samples, changes in relative concentrations for malabsorptive (MAL) and restrictive (RES) groups are plotted at each time point. The no bariatric surgery (NBS) group was used as the reference group in fold change calculations. *T1 11-14 weeks gestation; T2 20-24 weeks gestation; T3 28-30 weeks gestation; T4 30-33 weeks gestation; T5 35-37 weeks gestation; T6 delivery; GlycA N-acetyl glycoprotein; IS indoxyl sulfate; PAG phenylacetylglutamine; PCS p-cresol sulfate; PHPA p-hydroxyphenylacetate; *P_{adj} <0.05 (Mann-Whitney U)*

Table S5: Maternal metabolic changes depend on type of bariatric surgery. For each discriminatory metabolite identified by OPLS-DA (malabsorptive versus no bariatric surgery), changes in relative concentrations are detailed at each time point with no bariatric surgery as the reference group. Calculations in the restrictive group are shown for comparison.

Metabolite	Biofluid	Time point	Restrictive			Malabsorptive		
			log_2FC	*P*	P_{adj}	log_2FC	*P*	P_{adj}
Isoleucine	Serum	T1	0.036	NS	NS	-0.119	NS	NS
		T2	-0.149	3.15E-02	NS	-0.128	3.54E-02	NS
		T3	-0.053	NS	NS	-0.114	2.14E-02	4.18E-02
		T4	0.033	NS	NS	-0.156	1.11E-02	2.56E-02
		T5	-0.041	NS	NS	-0.168	4.62E-03	1.49E-02
		T6	-0.02	NS	NS	-0.061	NS	NS
Leucine	Serum	T1	-0.04	NS	NS	-0.093	NS	NS
		T2	-0.152	7.02E-03	NS	-0.105	2.65E-02	4.97E-02
		T3	-0.049	NS	NS	-0.062	NS	NS
		T4	-0.003	NS	NS	-0.133	2.39E-03	8.97E-03
		T5	0.077	NS	NS	-0.174	1.26E-04	1.14E-03
		T6	-0.028	NS	NS	-0.035	NS	NS
Isobutyrate	Serum	T1	0.032	NS	NS	-0.184	1.45E-03	5.69E-03
		T2	-0.077	NS	NS	-0.116	NS	NS
		T3	-0.044	NS	NS	-0.203	3.61E-04	1.86E-03
		T4	0.026	NS	NS	-0.191	2.54E-04	1.76E-03
		T5	0.062	NS	NS	-0.182	3.70E-04	1.86E-03
		T6	0.137	NS	NS	-0.014	NS	NS
Glutamine	Serum	T1	0.038	NS	NS	-0.008	NS	NS
		T2	0.009	NS	NS	0.031	NS	NS
		T3	-0.046	NS	NS	0.016	NS	NS
		T4	-0.032	NS	NS	0.044	NS	NS
		T5	-0.037	NS	NS	0.145	4.81E-02	NS
		T6	-0.242	NS	NS	-0.224	NS	NS
D-β-hydroxybutyrate	Serum	T1	0.112	NS	NS	-0.026	NS	NS
		T2	-0.096	NS	NS	-0.018	NS	NS
		T3	0.056	NS	NS	0.047	NS	NS
		T4	0.1	NS	NS	0.03	NS	NS
		T5	0.07	NS	NS	0.167	NS	NS
		T6	0.008	NS	NS	0.847	NS	NS
GlycA	Serum	T1	0.007	NS	NS	0.014	NS	NS
		T2	-0.007	NS	NS	-0.046	NS	NS
		T3	-0.039	NS	NS	-0.046	NS	NS
		T4	-0.062	NS	NS	-0.105	1.57E-02	3.37E-02
		T5	-0.079	NS	NS	-0.099	2.02E-02	4.04E-02
		T6	0.067	NS	NS	-0.04	NS	NS
PAG	Urine	T1	0.082	NS	NS	0.85	1.02E-04	1.02E-03
		T2	0.469	NS	NS	0.854	2.99E-05	4.38E-04
		T3	0.137	NS	NS	0.853	4.73E-05	5.32E-04
		T4	0.397	NS	NS	0.794	3.39E-05	4.38E-04
		T5	0.222	NS	NS	0.97	1.27E-06	1.14E-04
		T6	-0.064	NS	NS	0.162	NS	NS
		T1	0.178	NS	NS	0.631	1.42E-03	5.69E-03
		T2	0.375	NS	NS	0.856	2.40E-04	1.76E-03

PCS	Urine	T3	0.028	NS	NS	0.547	1.49E-02	3.26E-02
		T4	0.228	NS	NS	0.444	1.01E-02	2.39E-02
		T5	0.13	NS	NS	0.581	3.69E-03	1.28E-02
		T6	0.04	NS	NS	0.286	NS	NS
IS	Urine	T1	0.151	NS	NS	0.505	4.10E-02	NS
		T2	0.035	NS	NS	0.365	8.85E-03	2.28E-02
		T3	-0.081	NS	NS	0.478	8.02E-03	2.12E-02
		T4	0.069	NS	NS	0.388	6.51E-03	1.77E-02
		T5	0.212	NS	NS	0.739	1.83E-05	4.38E-04
		T6	0.105	NS	NS	0.297	NS	NS
PHPA	Urine	T1	-0.077	NS	NS	0.305	7.77E-04	3.50E-03
		T2	0.854	4.51E-02	NS	0.47	1.34E-03	5.69E-03
		T3	0.905	2.53E-02	NS	0.539	3.41E-05	4.38E-04
		T4	0.764	4.42E-03	NS	0.3	9.69E-03	2.39E-02
		T5	0.279	3.34E-02	NS	0.296	4.98E-03	1.55E-02
		T6	0.767	NS	NS	0.321	4.20E-02	NS
Unknown	Urine	T1	2.495	8.54E-03	NS	0.463	3.60E-04	1.86E-03
		T2	1.025	4.19E-02	NS	0.459	6.08E-03	1.73E-02
		T3	0.1	NS	NS	1.111	6.15E-03	1.73E-02
		T4	0.375	3.36E-02	NS	1.252	2.68E-05	4.38E-04
		T5	1.319	1.89E-03	NS	1.055	3.71E-04	1.86E-03
		T6	1.021	4.69E-02	NS	-0.343	NS	NS
Creatinine	Urine	T1	-0.282	NS	NS	-0.505	3.10E-05	4.38E-04
		T2	-0.057	NS	NS	-0.333	2.53E-03	9.12E-03
		T3	-0.096	NS	NS	-0.119	NS	NS
		T4	-0.159	NS	NS	-0.055	NS	NS
		T5	-0.045	NS	NS	-0.086	NS	NS
		T6	0.141	NS	NS	-0.173	NS	NS
α-ketoisovalerate	Urine	T1	-0.327	NS	NS	-0.504	2.24E-04	1.76E-03
		T2	-0.161	NS	NS	-0.284	NS	NS
		T3	-0.078	NS	NS	-0.307	1.77E-02	3.63E-02
		T4	-0.087	NS	NS	-0.301	2.27E-02	4.36E-02
		T5	0.073	NS	NS	-0.335	1.77E-02	3.63E-02
		T6	0.106	NS	NS	-0.23	NS	NS
Methylmalonate	Urine	T1	-0.224	NS	NS	-0.242	3.87E-02	NS
		T2	-0.191	NS	NS	-0.143	1.27E-02	2.86E-02
		T3	-0.133	NS	NS	-0.23	3.83E-03	1.28E-02
		T4	-0.148	NS	NS	-0.153	NS	NS
		T5	-0.073	NS	NS	-0.198	1.00E-02	2.39E-02
		T6	-0.102	NS	NS	0.038	NS	NS
Valine	Urine	T1	-0.13	NS	NS	-0.204	5.31E-03	1.59E-02
		T2	-0.175	8.95E-03	NS	-0.204	3.72E-04	1.86E-03
		T3	0.021	2.32E-02	NS	0.097	NS	NS
		T4	-0.071	NS	NS	-0.1	3.67E-02	NS
		T5	-0.169	9.34E-03	NS	-0.224	4.80E-04	2.28E-03
		T6	-0.237	NS	NS	-1.649	NS	NS

GlycA N-acetyl glycoprotein; PAG phenylacetylglutamine; PCS p-cresol sulfate; IS indoxyl sulfate; PHPA p-hydroxyphenylacetate; T1 11-14 weeks gestation; T2 20-24 weeks gestation; T3 28-30 weeks gestation; T4 30-33 weeks gestation; T5 35-37 weeks gestation; T6 delivery; FC fold change; NS not significant (P<0.05)

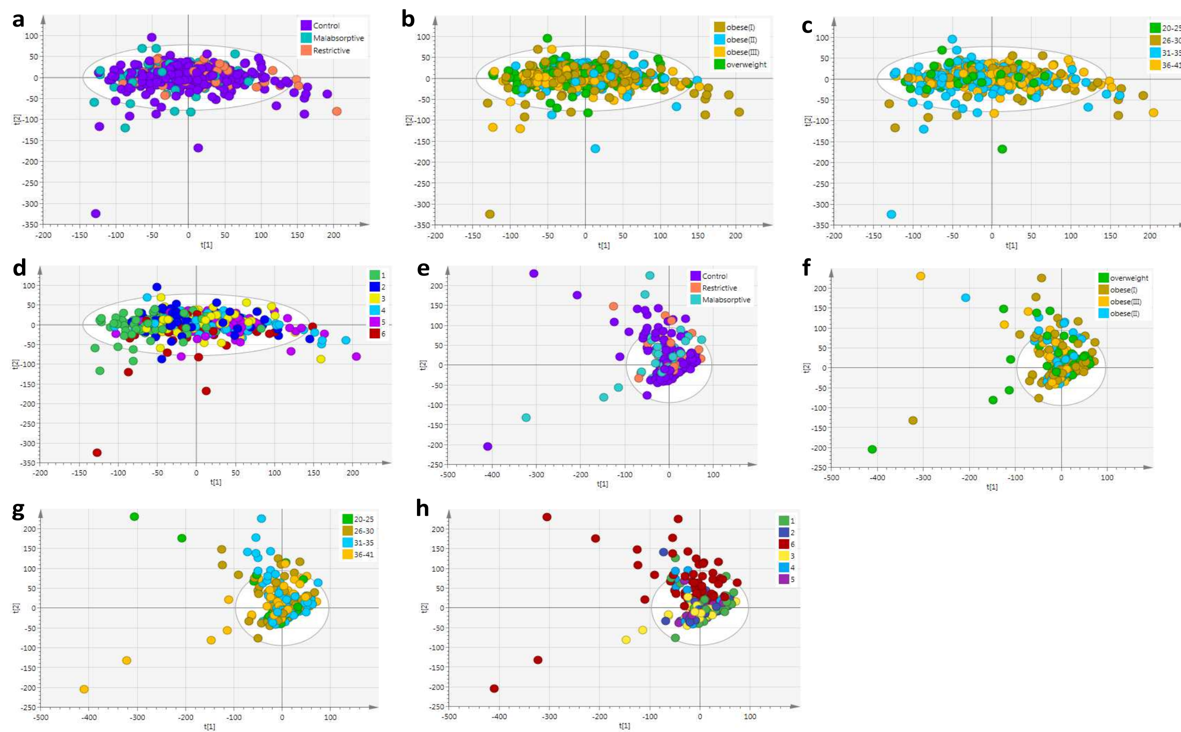

Figure S4: Principal component analysis (PCA) of ^{1}H NMR spectral data. Unsupervised PCA models were calculated with all samples analysed to assess overall variation in metabolic profiles. Serum metabolic profiles did not vary according to a) study group, b) BMI, or c) age but variation due to d) time point was observed between first trimester (T1) samples and those taken during later pregnancy (T2-T5) or at delivery (T6). Urinary metabolic profiles did not display overt signatures according to e) study group, f) BMI, or g) age but variation but due to h) time point was evident between samples taken during pregnancy (T1-T5) and those taken at delivery (T6). *T1 11-14 weeks gestation; T2 20-24 weeks gestation; T3 28-30 weeks gestation; T4 30-33 weeks gestation; T5 35-37 weeks gestation; T6 delivery*

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

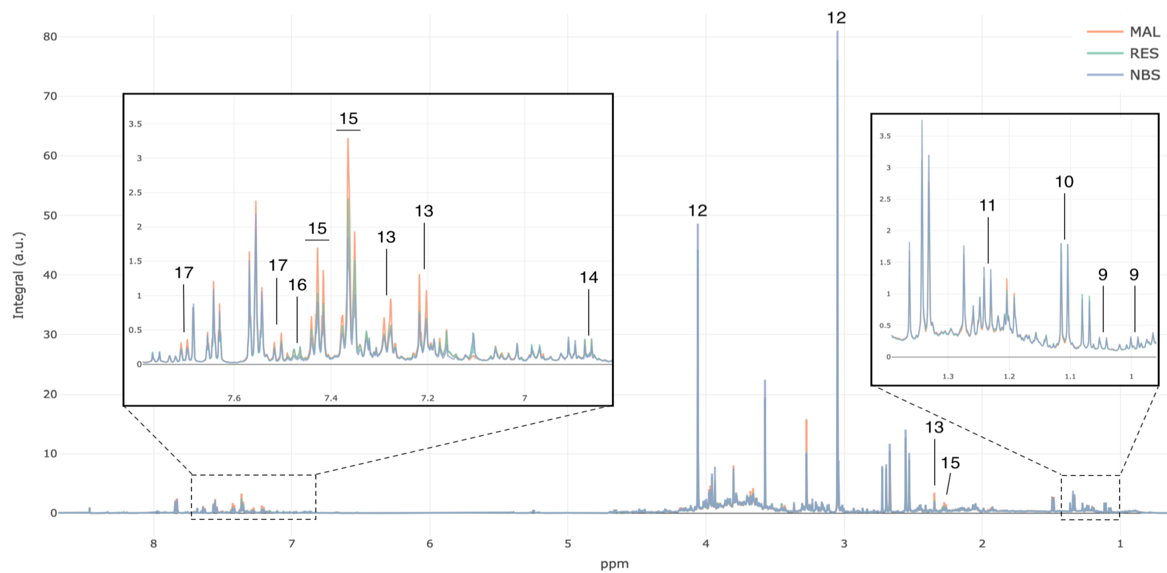

Figure S5: Overlayed median ^{1}H NMR urine spectra. A median spectrum is plotted for each study group with annotated peaks and peak groups. Annotated metabolites were considered discriminatory in OPLS-DA models comparing malabsorptive (MAL) and no bariatric surgery (NBS) patients. The representative spectrum for restrictive (RES) patients is also plotted. 9 valine; 10 α-ketoisovalerate; 11 methylmalonate; 12 creatinine; 13 *p*-cresol sulfate; 14 *p*-hydroxyphenylacetate; 15 phenylacetylglutamine; 16 unknown; 17 indoxyl sulfate

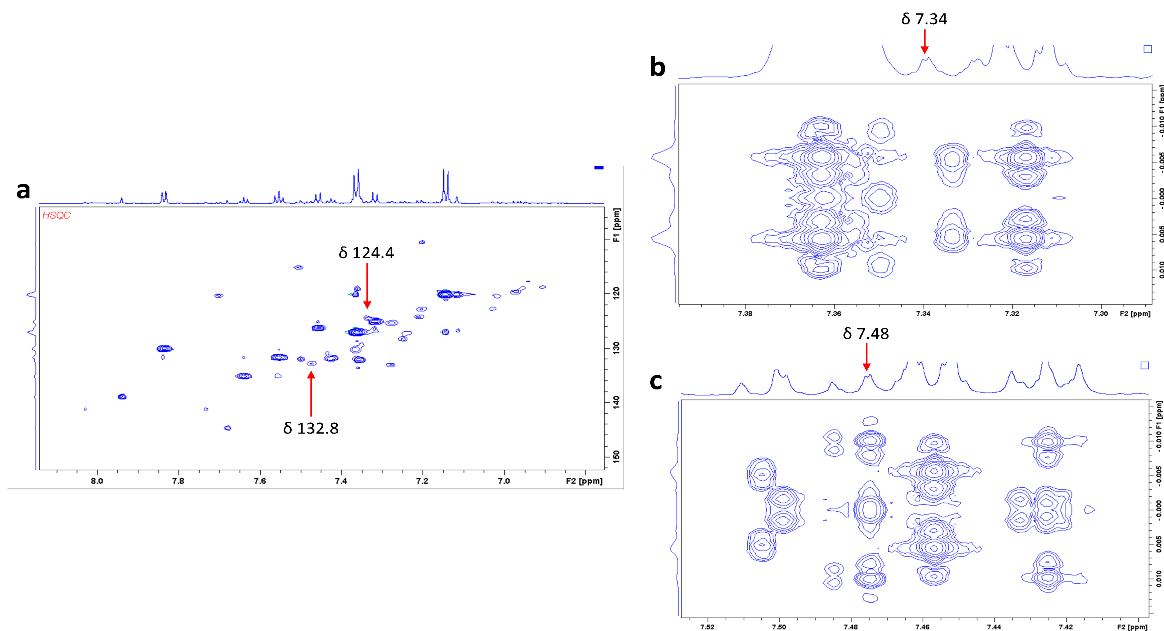

Figure S6: Chemical shifts of unknown metabolite. a) Hetero-nuclear Single Quantum Coherence (HSQC) spectrum detailing ^{13}C chemical shifts and b-c) J-resolved spectra detailing the proton-proton coupling of the ^{1}H chemical shifts.

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

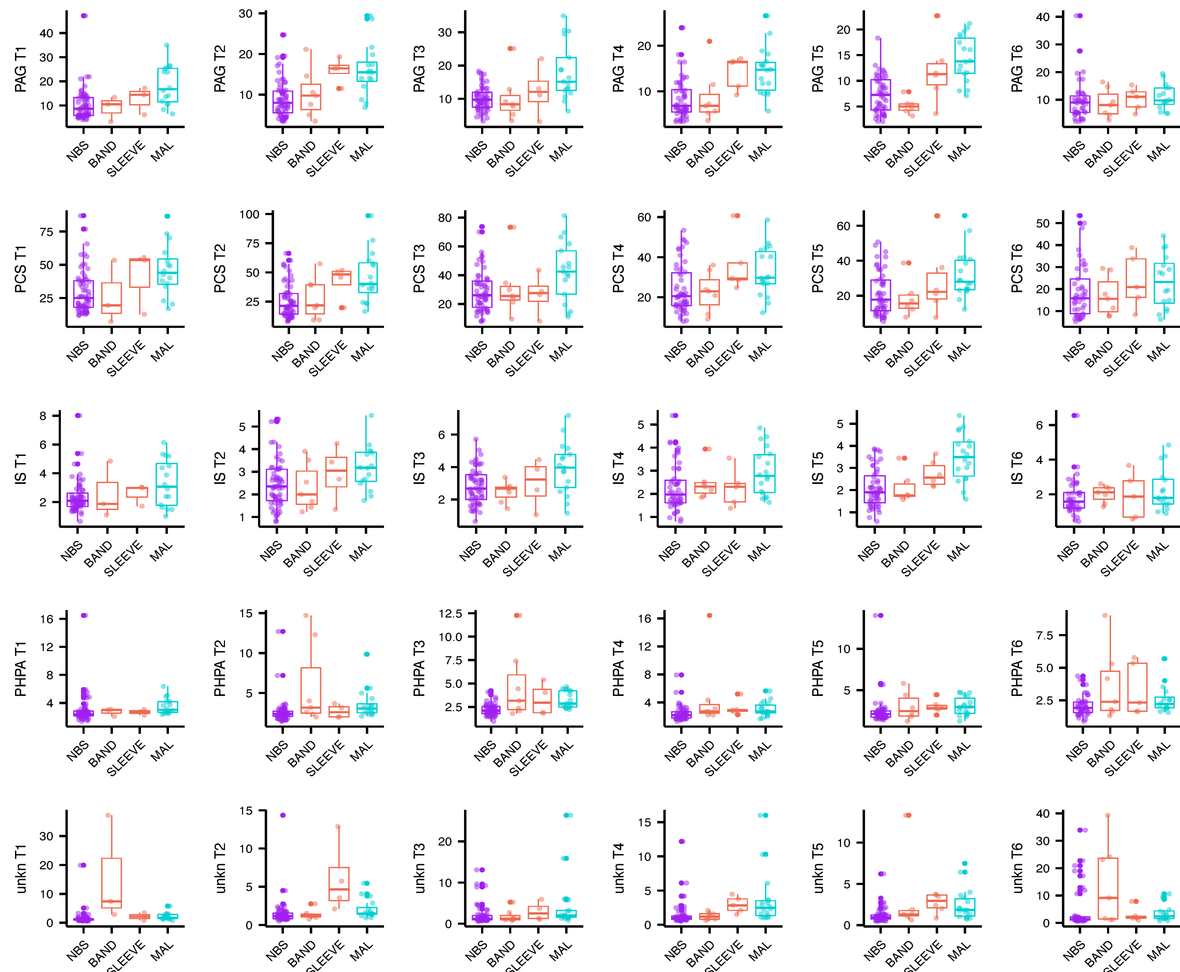

Figure S7: Relative concentrations of urinary host-microbial co-metabolites vary depending on bariatric surgery sub-type. Relative concentrations of each metabolite (by row) at each time point (by column) are represented in arbitraty units. Restrictive patients are separated by surgery sub-type (band or sleeve). *NBS no bariatric surgery; MAL malabsorptive; PAG phenylacetylglutamine; PCS p-cresol sulfate; IS indoxyl sulfate; PHPA p-hydroxyphenylacetate; unkn unknown ; T1 11-14 weeks gestation; T2 20-24 weeks gestation; T3 28-30 weeks gestation; T4 30-33 weeks gestation; T5 35-37 weeks gestation; T6 delivery*

Table S6: Differentially abundant taxa in malabsorptive group. For each taxonomic rank, differential abundance was assessed by calculating log_2(FC) of taxa relative abundances at each time point and applying the Mann-Whitney U test. Taxa with P_{adj} *<0.05* are reported.

Rank	Time point	Taxon	Mean relative abundance malabsorptive	Mean relative abundance control	*P* value	log_2(FC)	P_{adj}
Phylum	T1	*Proteobacteria*	0.13632	0.02701	0.0017	2.34	0.0237
Phylum	T4	*Bacteroidetes*	0.39265	0.52280	0.0016	-0.41	0.0237
Phylum	T4	*Proteobacteria*	0.16318	0.03249	5.45E-05	2.33	0.0020
Class	T1	*Gammaproteobacteria*	0.12379	0.01714	0.0022	2.85	0.0248
Class	T1	*Clostridia*	0.28025	0.40068	0.0048	-0.52	0.0357
Class	T1	*Bacilli*	0.05978	0.00789	0.0007	2.92	0.0144
Class	T2	*Bacilli*	0.03774	0.00694	0.0011	2.44	0.0202
Class	T4	*Bacteroidia*	0.39459	0.52538	0.0019	-0.41	0.0248
Class	T4	*Gammaproteobacteria*	0.13910	0.01069	2.61E-05	3.70	0.0020
Class	T4	*Bacilli*	0.07171	0.00987	5.45E-05	2.86	0.0020
Order	T1	*Enterobacteriales*	0.12097	0.01627	0.0048	2.89	0.0357
Order	T1	*Clostridiales*	0.28024	0.40059	0.0048	-0.52	0.0357
Order	T1	*Lactobacillales*	0.05906	0.00789	0.0007	2.90	0.0144
Order	T2	*Enterobacteriales*	0.11434	0.02743	0.0066	2.06	0.0425
Order	T2	*Lactobacillales*	0.03772	0.00689	0.0011	2.45	0.0202
Order	T4	*Bacteroidales*	0.39470	0.52539	0.0019	-0.41	0.0248
Order	T4	*Enterobacteriales*	0.12698	0.00888	4.54E-05	3.84	0.0020
Order	T4	*Lactobacillales*	0.07137	0.00982	5.45E-05	2.86	0.0020
Order	T4	*Pasteurellales*	0.01194	0.00181	0.0043	2.72	0.0357
Order	T4	*Actinomycetales*	0.00084	0.00025	0.0043	1.76	0.0357
Family	T1	*Enterobacteriaceae*	0.12514	0.01639	0.0048	2.93	0.0357
Family	T1	*Streptococcaceae*	0.05582	0.00795	0.0029	2.81	0.0285
Family	T1	*Micrococcaceae*	0.00028	8.76E-05	0.0042	1.66	0.0357
Family	T2	*Enterobacteriaceae*	0.11799	0.02810	0.0066	2.07	0.0425
Family	T2	*Enterococcaceae*	0.00290	1.30E-05	0.0003	7.80	0.0086
Family	T2	*Streptococcaceae*	0.03420	0.00689	0.0021	2.31	0.0248
Family	T2	*Micrococcaceae*	0.00027	4.95E-05	0.0028	2.45	0.0285
Family	T4	*Enterobacteriaceae*	0.12905	0.00927	4.54E-05	3.80	0.0020
Family	T4	*Enterococcaceae*	0.00179	2.78E-05	0.0003	6.01	0.0086
Family	T4	*Streptococcaceae*	0.07027	0.00983	5.45E-05	2.84	0.0020
Family	T4	*Pasteurellaceae*	0.01203	0.00183	0.0043	2.72	0.0357
Family	T4	*Carnobacteriaceae*	0.00032	6.41E-05	0.0013	2.30	0.0222
Family	T4	*Micrococcaceae*	0.00023	8.65E-05	0.0079	1.40	0.0497
Genus	T1	*Escherichia.Shigella*	0.11566	0.01274	0.0022	3.18	0.0248

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

Genus	T1	*Streptococcus*	0.06395	0.00895	0.0022	2.84	0.0248
Genus	T1	*Anaerostipes*	0.00281	0.01592	0.0037	-2.50	0.0350
Genus	T1	*Holdemanella*	0.02151	0.00238	0.0066	3.18	0.0425
Genus	T1	*Rothia*	0.00031	9.97E-05	0.0055	1.64	0.0380
Genus	T2	*Enterococcus*	0.00323	1.45E-05	0.0003	7.80	0.0086
Genus	T2	*Streptococcus*	0.03881	0.00794	0.0028	2.29	0.0285
Genus	T2	*Butyricimonas*	0.00191	0.00029	0.0050	2.70	0.0357
Genus	T2	*Rothia*	0.00030	5.59E-05	0.0028	2.43	0.0285
Genus	T4	*Escherichia.Shigella*	0.12798	0.00838	5.44E-05	3.93	0.0020
Genus	T4	*Enterococcus*	0.00202	3.12E-05	0.0003	6.02	0.0086
Genus	T4	*Streptococcus*	0.07909	0.01096	5.45E-05	2.85	0.0020
Genus	T4	*Anaerostipes*	0.00395	0.01148	0.0064	-1.54	0.0425
Genus	T4	*Haemophilus*	0.01321	0.00202	0.0049	2.71	0.0357
Genus	T4	*Veillonella*	0.00982	0.00120	0.0004	3.04	0.0097
Genus	T4	*Acidaminococcus*	0.00350	7.71E-05	0.0044	5.50	0.0357
Genus	T4	*Solobacterium*	0.00118	0.00029	0.0014	2.01	0.0223
Genus	T4	*Granulicatella*	0.00035	7.13E-05	0.0015	2.31	0.0233

FC fold change; T1 11-14 weeks gestation; T2 20-24 weeks gestation; T4 30-33 weeks gestation

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

Table S7: Differentially abundant amplicon sequence variants. 16S rRNA gene amplicon sequences belonging to genera of interest that are significantly ($P_{adj}<0.05$) more abundant (DESeq2) in the malabsorptive group compared to no bariatric surgery controls.

Genus	log_2(FC)	P_{adj}	Species assignment	
			RDP	SILVA
Escherichia/ Shigella	4.96	0.0132	*coli*	*coli/flexneri*
	9.31	0.0005	-	*coli*
	7.15	2.40E-05	*coli*	*coli/fergusonii*
	3.95	0.0412	*coli/dysenteriae*	-
	5.71	0.0357	-	*coli*
	7.3	1.20E-05	*coli*	*coli*
	5.21	0.0293	*coli/flexneri*	*coli/flexneri/sonnei*
	3.3	0.0184	-	-
	22.69	6.80E-12	*coli*	*coli*
	6.27	3.60E-06	*coli*	*coli/fergusonii*
	3.31	0.0002	-	-
Streptococcus	5.26	0.0408	*gordonii*	*gordonii/sanguinis*
	7.52	1.10E-13	*vestibularis*	*vestibularis*
	3.68	6.70E-07	*salivarius/ vestibularis*	-
	4.09	0.0076	*salivarius/ vestibularis*	*salivarius*
	5.4	0.0376	*salivarius*	*salivarius*
	4.24	1.40E-05	-	*infantis*
	3.45	0.0313	*mitis/sanguinis*	-
	4.05	0.0001	*mutans*	*mutans*
	3.61	0.0297	*parasanguinis*	*parasanguinis*
	3.39	0.0201	*mitis/parasanguinis*	*mitis*
	2.78	0.0135	*parasanguinis*	*parasanguinis*
	3.34	0.0049	-	*parasanguinis*
Enterococcus	6.98	1.00E-12	*faecalis/faecium*	*faecalis*
Rothia	2.01	0.0316	*mucilaginosa*	-

FC fold change

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

Table S8: Relative concentrations of metabolites correlate with maternal insulin resistance and fetal/birth weight. Spearman's (ρ) correlations between metabolite relative concentrations at each time point and clinical or biochemical measures. Statistically significant (P_{adj} *<0.05)* correlations are detailed along with partial correlations.

Variable 1	Variable 2	ρ	*P*	P_{adj}	Partial ρ	*P*	P_{adj}
PAG_T1	Fasting Maternal Insulin (microU/ml)	-0.37	0.00063	0.00751	-0.28	0.01335	0.02114
PCS_T1	Fasting Maternal Insulin (microU/ml)	-0.36	0.00111	0.00974	-0.28	0.01229	0.02031
IS_T1	Fasting Maternal Insulin (microU/ml)	-0.31	0.00447	0.02349	-0.30	0.00753	0.01607
PAG_T2	Fasting Maternal Insulin (microU/ml)	-0.46	0.00001	0.00040	-0.32	0.00327	0.00879
PCS_T2	Fasting Maternal Insulin (microU/ml)	-0.35	0.00076	0.00796	-0.23	0.03351	0.03979
Ile_T2	Fasting Maternal Insulin (microU/ml)	0.4	0.00052	0.00727	0.38	0.00119	0.00664
Leu_T2	Fasting Maternal Insulin (microU/ml)	0.38	0.00116	0.00974	0.37	0.00184	0.00865
PAG_T3	Fasting Maternal Insulin (microU/ml)	-0.46	0.00001	0.00044	-0.33	0.00269	0.00865
PCS_T3	Fasting Maternal Insulin (microU/ml)	-0.42	0.00006	0.00155	-0.33	0.00252	0.00865
PAG_T4	Fasting Maternal Insulin (microU/ml)	-0.4	0.00021	0.00348	-0.36	0.00101	0.00664
PCS_T4	Fasting Maternal Insulin (microU/ml)	-0.4	0.00018	0.00328	-0.39	0.00041	0.00466
PAG_T5	Fasting Maternal Insulin (microU/ml)	-0.48	0.00001	0.00040	-0.39	0.00049	0.00466
PCS_T5	Fasting Maternal Insulin (microU/ml)	-0.41	0.00023	0.00353	-0.30	0.00858	0.01717
IS_T5	Fasting Maternal Insulin (microU/ml)	-0.41	0.00017	0.00328	-0.36	0.00122	0.00664
Leu_T5	Fasting Maternal Insulin (microU/ml)	0.36	0.00444	0.02349	0.34	0.00761	0.01607
PCS_T6	Fasting Maternal Insulin (microU/ml)	-0.32	0.00719	0.03551	-0.22	NS	NS
PAG_T1	Maternal HOMA-IR	-0.35	0.00148	0.01068	-0.25	0.02701	0.03311
PCS_T1	Maternal HOMA-IR	-0.35	0.00161	0.01079	-0.27	0.01638	0.02305
IS_T1	Maternal HOMA-IR	-0.29	0.00822	0.03943	-0.27	0.01488	0.02262
PAG_T2	Maternal HOMA-IR	-0.46	0.00001	0.00040	-0.32	0.00347	0.00879
PCS_T2	Maternal HOMA-IR	-0.33	0.00153	0.01068	-0.20	NS	NS
Ile_T2	Maternal HOMA-IR	0.38	0.00089	0.00828	0.35	0.00261	0.00865
Leu_T2	Maternal HOMA-IR	0.33	0.00418	0.02341	0.31	0.00994	0.01808
PAG_T3	Maternal HOMA-IR	-0.41	0.00010	0.00235	-0.26	0.01813	0.02461
PCS_T3	Maternal HOMA-IR	-0.35	0.00081	0.00796	-0.23	0.03463	0.03988

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

Ile_T3	Maternal HOMA-IR	0.3	0.00991	0.04500	0.20	NS	NS
PAG_T4	Maternal HOMA-IR	-0.37	0.00068	0.00766	-0.32	0.00340	0.00879
PCS_T4	Maternal HOMA-IR	-0.35	0.00148	0.01068	-0.31	0.00470	0.01115
PAG_T5	Maternal HOMA-IR	-0.44	0.00005	0.00155	-0.34	0.00273	0.00865
PCS_T5	Maternal HOMA-IR	-0.33	0.00298	0.01857	-0.21	NS	NS
IS_T5	Maternal HOMA-IR	-0.36	0.00138	0.01068	-0.29	0.01047	0.01808
Leu_T5	Maternal HOMA-IR	0.32	0.01084	0.04794	0.29	0.02488	0.03260
PAG_T2	EFW T2	-0.31	0.00321	0.01925	-0.26	0.01582	0.02305
PCS_T2	EFW T2	-0.29	0.00657	0.03344	-0.28	0.01014	0.01808
Ile_T2	EFW T2	0.3	0.00949	0.04428	0.20	NS	NS
IS_T5	EFW T5	-0.32	0.00402	0.02328	-0.26	0.02681	0.03311
PAG_T6	Birth weight	-0.36	0.00270	0.01743	-0.46	0.00020	0.00371
IS_T6	Birth weight	-0.4	0.00059	0.00751	-0.49	5.22E-05	0.00198351

PAG phenylacetylglutamine; PCS p-cresol sulfate; IS indoxyl sulfate; Ile isoleucine; Leu leucine; HOMA-IR Homeostatic Model Assessment of Insulin Resistance; EFW estimated fetal weight; T1 11-14 weeks gestation; T2 20-24 weeks gestation; T3 28-30 weeks gestation; T4 30-33 weeks gestation; T5 35-37 weeks gestation; T6 delivery

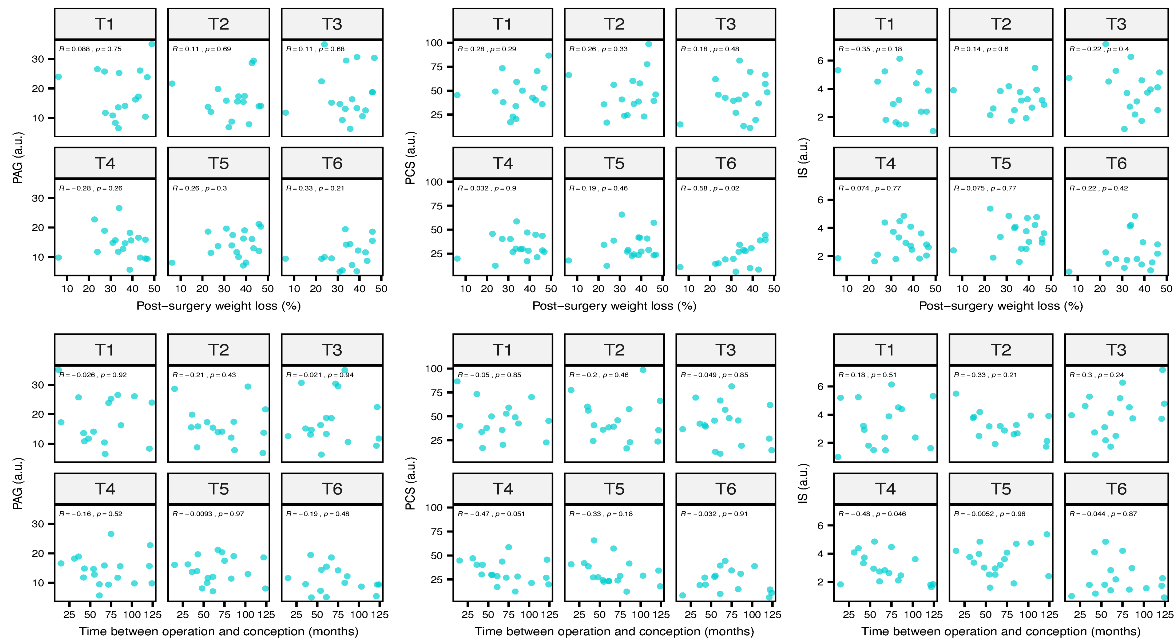

Figure S8: Key metabolites are not associated with percentage of weight lost or length of time since surgery in malabsorptive patients. Spearman correlations between metabolite relative concentrations at each time point and (a) percentage of weight lost from time of surgery to conception or (b) number of months between time of surgery and conception. Relative concentrations are reported in arbitrary units (a.u.). *PAG phenylacetylglutamine; PCS p-cresol sulfate; IS indoxyl sulfate.*

SUPPLEMENTARY METHODS

Longitudinal metabolic and gut bacterial profiling of pregnant women with previous bariatric surgery

Study population, sampling and clinical data

The population is part of an ongoing prospective study investigating the impact of maternal BS on perinatal outcomes. The study was approved by the West London Research Ethics Committee (No: 14/LO/0592) and all women gave written, informed consent for their data and samples to be used. Pregnant women with and without previous BS were recruited from May 2015 to April 2017 at Chelsea & Westminster Hospital (London, UK) as previously described[1]. Women were seen at 5 time points during pregnancy (**T1**: 11^{+0}-14^{+0}, **T2**: 20^{+0}-24^{+0}, **T3**: 28^{+0}-30^{+0}, **T4**: 30^{+0}-33^{+0} and **T5**: 35^{+0}-37^{+6} weeks gestation) and within 72 h of delivery (**T6**). Maternal blood (serum) and urine samples were collected at each visit while fecal samples were requested at the T1, T2 and T4 visits (see online supplementary figure S1 and table S1). A full oral glucose tolerance test (2 h, 75 g) was conducted at T3 and maternal insulin resistance was calculated using the homeostatic model assessment for insulin resistance (HOMA-IR=fasting serum insulin (µU/L) x fasting glucose (mmol/L)/22.5)[2]. Estimated fetal weight was calculated by trans-abdominal ultrasound scans[3] at T2, T4 and T5. At T6, birth weight was recorded, percentiles for the gestation were calculated[4] and, where possible, neonatal samples (cord serum and urine) were collected. All samples were stored at -80°C for future analysis. For the current study population, women with diagnosis of type 2 diabetes mellitus or GDM (due to the effect of diabetes on the metabolic profile) and those that had a miscarriage were excluded. Only NBS participants with a BMI of 25 to 50 kg/m^2 at T1 were included to match the BMI range of the included bariatric patients at T1.

Metabolic profiling of biofluid samples

Serum and urine samples were prepared according to an established protocol[5]. ^{1}H NMR spectra were acquired on a Bruker 600 MHz spectrometer (Bruker BioSpin, Karlsruhe, Germany) following a

published method[5]. Briefly, experiments for serum samples were run at 310 K using a Carr-Purcell-Meiboom-Gill (CPMG) spin-echo pulse sequence with water presaturation (Bruker pulse program: cpmgpr1d). For urinary samples, experiments were run at 300 K using a standard one-dimensional (1D) spectroscopy pulse sequence with water presaturation (noesygppr1d). J-resolved (J-res) experiments (jresgpprqf) were also run for all samples to increase capacity for structural identification of molecules. The spectra were automatically baseline corrected, phased and referenced to sodium 3-(trimethylsilyl) propionate-2,2,3,3-d_4 (TSP) in the Bruker TopSpin 3.1 software. Raw data were processed in MATLAB version R2016b (The MathWorks, Inc., Natwick, MA) using scripts developed by Dr. T. Ebbels at Imperial College. Spectra were aligned using the recursive segment-wise peak alignment (RSPA) algorithm[6] and normalised with the probabilistic quotient normalisation (PQN) function[7]. Downstream multivariate modelling was performed in SIMCA 14.1 (Sartorius Stedim Biotech, Aubagne, France).

Statistical modelling and analysis of metabolic profiles

Unsupervised principal component analysis (PCA) models were used to assess variation in metabolic profile over all time points and to identify extreme outliers (based on Hotelling's T^2 statistic) to exclude from supervised models for each time point. Orthogonal partial least squares discriminant analysis (OPLS-DA)[8] was used to identify spectral variables that contributed to discrimination of clinical classes (pairwise comparisons between NBS, MAL and RES groups) at each time point. Valid models were determined by positive Q^2Y value and significant (P<0.05) ANOVA of the cross-validated residuals (CV-ANOVA[9]). NMR peaks were considered discriminatory if their correlation with the predictive component was 0.45 or greater. Relative concentrations of metabolites were calculated by integrating a representative peak of each identified metabolite. Downstream analysis was performed in the R software environment[10]. Time-series curves were generated for each discriminatory metabolite with "santaR". In this method, the measurements for each individual are condensed into a smooth, continuous function of time and mean curves are then calculated for each study group. Individuals with at least 5 data points were included. Significance of the distance between group mean

trajectories is tested using a permutation-based method. Changes in metabolite concentrations at each time point were represented by log_2(Fold Change) where Fold Change (FC) = mean(Case)/mean(Control) and significance was assessed by Mann-Whitney U test. Spearman's correlation coefficients (ρ) and *P* values were calculated for metabolite-clinical associations with the package "Hmisc". Partial correlations were calculated with "ppcor"[11] to adjust for confounding variables (maternal age and BMI on maternal measurements and maternal age, BMI and HOMA-IR on birthweight percentile). Biospecimens with missing data were removed from the analysis. All *P* values were adjusted (P_{adj}) where necessary to control the false discovery rate according to the Benjamini-Hochberg method[12]. An alpha of 0.05 was used for *P* and P_{adj} values. Plots were generated with "ggpubr" apart from the correlation network visualisation which was generated with Cytoscape version 3.5.1[13].

Metabolite identification

Discriminatory NMR peaks were used as driver peaks for statistical total correlation spectroscopy (STOCSY)[14] and subset optimization by reference matching (STORM)[15], two algorithms used to identify additional peaks corresponding to the same molecule as the driver peak. Chemical shifts, multiplicities, and J-couplings for each molecule were matched to an in-house reference database for annotation. Ambiguous annotations were confirmed with 2D NMR experiments (J-res, HSQC[16]) and by spiking in authentic standards[5].

Gut bacterial community profiling

Stool samples were randomised for processing and DNA was extracted from 250 mg stool using the PowerLyzer PowerSoil DNA Isolation Kit (Mo Bio, Carlsbad, CA, USA). Bead beating was carried out in a Bullet Blender Storm (Chembio Ltd., St. Albans, UK) for 3 min at speed 8. DNA was quantified using the Qubit fluorometric assay (Thermo Fisher Scientific, Carlsbad, CA, USA). Sample libraries amplifying the V1-V2 region of the 16S rRNA gene were prepared as previously described[17] and were sequenced on the Illumina MiSeq platform (San Diego, CA, USA) using the MiSeq Reagent Kit v3 and

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

paired-end 300 bp chemistry. Primer sequences were removed from demultiplexed fastq files using cutadapt[18] and raw reads were processed in the R software environment[10] following a published workflow[19] which includes amplicon denoising[20]. The denoising algorithm "DADA2" uses error profiles generated during the sequencing run to infer real sequence variants; it allows the analysis of unique sequences rather than operational taxonomic units (OTUs), or clusters of similar sequences. Taxonomy was assigned with reference to the RDP[21] database and assignments of statistically significant taxa were confirmed with the SILVA[22] database.

Statistical analysis of microbiome data

Functions in the "vegan" R package were used to calculate Shannon Diversity Indices[23] on data rarefied to the minimum sequencing depth and Bray-Curtis dissimilarity[24] on log-transformed (pseudocount of 1 added to each value) data. Permutational multivariate analysis of variance (PERMANOVA)[25] was applied to the Bray-Curtis matrix to test whether the taxa distributions were different between the clinical classes. Changes in relative abundance were tested at each taxonomic rank from phylum to genus using the Mann-Whitney U test while differentially abundant 16S rRNA gene sequences were identified using "DESeq2" on raw counts[26]. DESeq2 implements a statistical model to account for the sparsity of the count matrix and over-dispersion of the counts, two features that are characteristic of 16S rRNA gene amplicon sequencing data. For "DESeq2" analysis, data were pooled for each individual rather than analysing distinct time points.

Integrative analysis of metabolic and taxonomic data

Relationships between the serum, urine and faecal datasets were modelled using the DIABLO method in "mixOmics"[27]. This is a multi-block latent variable-based approach which aims to identify concordance between multiple datasets. Metabolites significant throughout the time course (serum: leucine, isoleucine, isobutyrate, D-β-hydroxybutyrate; urine: PAG, PCS, IS, PHPA, unknown, α-ketoisovalerate, creatinine) and a subset of bacterial genera (log-transformed; selected using the LASSO penalization method implemented in "mixOmics") were modelled. Sampling points for each

individual where matching microbiome and metabolite data were obtained (T1, T2 and T4) were included in the model.

All R code, packages and package versions used for data analysis can be found from page 7 of this document. The R Markdown file and data to reproduce the analysis are availale on GitHub (https://github.com/ka-west/PBS_manuscript).

REFERENCES

1 Maric T, Kanu C, Johnson MR, Savvidou MD. Maternal, neonatal insulin resistance and neonatal anthropometrics in pregnancies following bariatric surgery. Metabolism 2019;**97**:25-31.
2 Matthews DR, Hosker JR, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and fl-cell function from fasting plasma glucose and insulin concentrations in man Diabetologia 1985;**28** 412-9.
3 Maršál K, Persson P- H, Larsen T, Lilja H, Selbing A, Sultan B. Intrauterine growth curves based on ultrasonically estimated foetal weights. Acta Paediatrica 1996;**85**:843-8.
4 Poon LCY, Volpe N, Muto B, Syngelaki A, Nicolaides KH. Birthweight with Gestation and Maternal Characteristics in Live Births and Stillbirths. Fetal Diagnosis and Therapy 2012;**32**:156-65.
5 Dona AC, Jiménez B, Schäfer H, Humpfer E, Spraul M, Lewis MR*, et al.* Precision High-Throughput Proton NMR Spectroscopy of Human Urine, Serum, and Plasma for Large-Scale Metabolic Phenotyping. Analytical Chemistry 2014;**86**:9887-94.
6 Veselkov KA, Lindon JC, Ebbels TMD, Crockford D, Volynkin VV, Holmes E*, et al.* Recursive Segment-Wise Peak Alignment of Biological 1H NMR Spectra for Improved Metabolic Biomarker Recovery. Analytical Chemistry 2009;**81**:56-66.
7 Dieterle F, Ross A, Schlotterbeck G, Senn H. Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in 1H NMR Metabonomics. Analytical Chemistry 2006;**78**:4281-90.
8 Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. Journal of Chemometrics 2006;**20**:341-51.
9 Eriksson L, Trygg J, Wold S. CV-ANOVA for significance testing of PLS and OPLS® models. Journal of Chemometrics 2008;**22**:594-600.
10 R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. 2017.
11 Kim S. ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. Communications for statistical applications and methods 2015;**22**:665-74.
12 Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) 1995;**57**:289-300.
13 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D*, et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research 2003;**13**:2498-504.
14 Cloarec O, Dumas M-E, Craig A, Barton RH, Trygg J, Hudson J*, et al.* Statistical Total Correlation Spectroscopy: An Exploratory Approach for Latent Biomarker Identification from Metabolic 1H NMR Data Sets. Analytical Chemistry 2005;**77**:1282-9.
15 Posma JM, Garcia-Perez I, De Iorio M, Lindon JC, Elliott P, Holmes E*, et al.* Subset Optimization by Reference Matching (STORM): An Optimized Statistical Approach for Recovery of

Metabolic Biomarker Structural Information from 1H NMR Spectra of Biofluids. Analytical Chemistry 2012;**84**:10694-701.
16 Nicholson JK, Foxall PJD, Spraul M, Farrant RD, Lindon JC. 750 MHz 1H and 1H-13C NMR Spectroscopy of Human Blood Plasma. Analytical Chemistry 1995;**67**:793-811.
17 Mullish BH, Pechlivanis A, Barker GF, Thursz MR, Marchesi JR, McDonald JAK. Functional microbiomics: Evaluation of gut microbiota-bile acid metabolism interactions in health and disease. Methods (San Diego, Calif) 2018;**149**:49-58.
18 Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. Bioinformatics in Action 2012;**17**:10-2.
19 Callahan B, Sankaran K, Fukuyama J, McMurdie P, Holmes S. Bioconductor Workflow for Microbiome Data Analysis: from raw reads to community analyses [version 2; referees: 3 approved], 2016.
20 Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Meth 2016;**13**:581-3.
21 Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y*, et al.* Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Research 2014;**42**:D633-D42.
22 Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P*, et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 2013;**41**:D590-D6.
23 Shannon CE. A mathematical theory of communication. SIGMOBILE Mob Comput Commun Rev 2001;**5**:3-55.
24 Bray JR, Curtis JT. An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecological Monographs 1957;**27**:325-49.
25 Anderson M. A new method for non-parametric multivariate analysis of variance. Austral Ecology 2001;**26**:32-46.
26 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 2014;**15**:550.
27 Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: An R package for ‘omics feature selection and multiple data integration. PLOS Computational Biology 2017;**13**:e1005752.

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

R script - report generated with "knitr"

R Markdown file and data to reproduce this analysis can be found on GitHub at https://github.com/ka-west/PBS_manuscript

---

All code used to generate results within the manuscript and the supplementary files is contained here

- For all *P* values, an alpha of 0.05 was used
- The Benjamini-Hochberg method was used to control for false discovery rate

---

Parameters

```
save_dir <- "output" # where to save figures and tables
if (!dir.exists(save_dir)) { stop(paste("save_dir", save_dir, "does not exist")) }

color_pal <- list(NBS = "purple", RES = "coral2", MAL = "darkturquoise")
```

Load data

```
load("PBS_data.Rdata")
ls()
```

```
## [1] "color_pal"       "integrals"       "med_specs_serum" "med_specs_urine"
## [5] "pkgs"            "ps_M"            "rdp.species"     "save_dir"
## [9] "silva.species"
```

Figure S1: Summary of sampling by individual

Contribution of samples by study group at each time point

```
mb <- data.frame(Patient.TP = sample_names(ps_M),
                 faeces = TRUE)

# order individuals by number of time points where samples were collected
ord <- table(integrals$Study.no) %>%
  sort(decreasing = TRUE) %>%
  names()

# which samples were collected for each individual at each time point
all_samples <- integrals %>%
  filter(Time_point != 7) %>%
  mutate(urine = ifelse(!is.na(PAG), TRUE, NA),
         serum = ifelse(!is.na(isoleucine), TRUE, NA)) %>%
  full_join(mb, by = "Patient.TP") %>%
  mutate(Sampling = case_when(urine == TRUE & serum == TRUE & faeces == TRUE ~ "Serum/Urine/Faeces",
                              urine == TRUE & serum == TRUE & is.na(faeces) ~ "Serum/Urine",
                              urine == TRUE & is.na(serum) & faeces == TRUE ~ "Urine/Faeces",
                              is.na(urine) & serum == TRUE & faeces == TRUE ~ "Serum/Faeces",
                              urine == TRUE & is.na(serum) & is.na(faeces) ~ "Urine only",
                              is.na(urine) & serum == TRUE & is.na(faeces) ~ "Serum only"),
         Sampling = factor(Sampling, levels = c("Serum/Urine/Faeces", "Serum/Urine", "Urine/Faeces",
                                                "Serum/Faeces", "Urine only",  "Serum only")),
         Study.no = factor(Study.no, levels = ord))
```

Make plot

```
p <- ggscatter(all_samples,
               x = "Study.no",
               xlab = "Individual",
               y = "Time_point_label",
               ylab = "Time point",
               color = "Sampling",
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
              palette = c("chartreuse3", "mediumorchid3", "dodgerblue",
                          "darkorange", "gold2", "red3"),
              size = "Obesity") +
  theme(axis.text.x = element_blank(),
        axis.ticks.x = element_blank(),
        legend.spacing.x = unit(2, "mm"),
        legend.title = element_blank()) +
  grids(linetype = "dashed", axis = "x", color = "grey")

p <- facet(p, facet.by = "Group", scales = "free", ncol = 1)
```

Figures S2 and S5: Overlayed median 1H NMR spectra

Serum (Figure S2) and urine (Figure S5) median spectra were overlayed for the three study groups

```
plot_med_specs <- function(to_plot, save_dir, save_name) {

  p <- plot_ly(to_plot, x = ~ppm, y = ~MAL, name = "MAL", type = "scatter",
               mode = "lines", alpha = 0.8, color = color_pal$MAL) %>%
    add_trace(y = ~RES, name = "RES", mode = "lines", alpha = 0.8, color = color_pal$RES) %>%
    add_trace(y = ~NBS, name = "NBS", mode = "lines", alpha = 0.8, color = color_pal$NBS) %>%
    layout(xaxis = list(autorange = "reversed", size = 18),
           yaxis = list(title = "Integral (a.u.)", size = 18))

  htmlwidgets::saveWidget(p, file = file.path(getwd(), save_dir, save_name))
}

# serum
plot_med_specs(med_specs_serum, save_dir, "figureS2_median_spectra_serum.html")

# urine
plot_med_specs(med_specs_urine, save_dir, "figureS5_median_spectra_urine.html")
```

Figure 1: Longitudinal modelling of metabolite concentrations

Metabolites were considered discriminatory (malabsorptive vs control) if their NMR peaks were correlated (>=0.45) with the predictive component of a valid OPLS-DA model at any time point. A representative peak for each metabolite was integrated to calculate relative concentration.

The behavior of key discriminatory metabolites over time was assessed and compared between control and malabsorptive groups using `santaR` time-series analysis.

Abbreviations: PAG: phenylacetylglutamine, PCS: *p*-cresol sulfate, IS: indoxyl sulfate, PHPA: *p*-hydroxyphenylacetate, unkn: unknown, aKIV: a-ketoisovalerate, MM: methylmalonate, THB: D-B-hydroxybutyrate, NAGP: N-acetyl glycoprotein

```
################## DEFINED VARIABLES ###################

# remove restrictive group
input_data <- subset(integrals, integrals$Group != "Restrictive" & integrals$Time_point != 7)

# column names
metab_names <- c("leucine", "isoleucine", "isobutyrate", "NAGP", "glutamine", "THB")
tp_col <- "Time_point"
patient_col <- "Study.no"
group_col <- "Group"

tp_labs <- c("T1", "T2", "T3", "T4", "T5", "T6")

plot_titles <- c("Serum - Leucine",
                 "Serum - Isoleucine",
                 "Serum - Isobutyrate",
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
                  "Serum - GlycA",
                  "Serum - Glutamine",
                  "Serum - D-beta-hydroxubutyrate")

# see santaR vignette for choosing this value
degrees_freedom <- 5

# add custom locations (y axis) on plots for p values
pval_locs = c(27, 20, 2.1, 5.0, 8.0, 17)

############################################################
```

Serum metabolites

```
# run santaR time series analysis
sp <- santaR_auto_fit(inputData = input_data[,metab_names], ind = input_data[[patient_col]],
                      time = as.numeric(input_data[[tp_col]]), group = input_data[[group_col]],
                      df = degrees_freedom)

# generate a plot for each metabolite
for (metab in 1:length(metab_names)) {

  p <- santaR_plot(sp[[metab_names[metab]]],
                   showIndPoint=FALSE,
                   showIndCurve=FALSE,
                   xlab = NULL,
                   ylab = NULL,
                   colorVect = unlist(color_pal[-2]),
                   title = plot_titles[metab]) +
    scale_x_continuous(breaks = 1:length(tp_labs),
                       labels = tp_labs) +
    theme(plot.title = element_text(size = 6),
          legend.position = "None",
          axis.text = element_text(size = 5),
          panel.grid.major = element_blank(),
          panel.grid.minor = element_blank()) +
    # add p value
    annotate("text",
             size = 2,
             x = 5,
             y = pval_locs[metab],
             label = paste("P = ", round(sp[[metab_names[metab]]]$general$pval.dist, digits = 4)))

  assign(metab_names[metab], p)
}
```

Urine metabolites

```
################## DEFINED VARIABLES ##################

# remove restrictive group
input_data <- subset(integrals, integrals$Group != "Restrictive" & integrals$Time_point != 7)

# column names
metab_names <- c("PAG", "PCS", "IS", "PHPA", "unkn", "valine", "aKIV", "MM", "creatinine")
tp_col <- "Time_point"
patient_col <- "Study.no"
group_col <- "Group"

tp_labs <- c("T1", "T2", "T3", "T4", "T5", "T6")
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```r
plot_titles <- c("Urine - PAG",
                 "Urine - PCS",
                 "Urine - IS",
                 "Urine - PHPA",
                 "Urine - Unknown",
                 "Urine - Valine",
                 "Urine - alpha-Ketoisovalerate",
                 "Urine - Methylmalonate",
                 "Urine - Creatinine")

# see santaR vignette for choosing this value
degrees_freedom <- 5

# add custom locations (y axis) on plots for p values
pval_locs = c(28, 75, 6.5, 6.5, 13, 12, 22, 12, 550)

###########################################################
```

```r
# run santaR time series analysis
sp <- santaR_auto_fit(inputData = input_data[,metab_names], ind = input_data[[patient_col]],
                      time = as.numeric(input_data[[tp_col]]), group = input_data[[group_col]],
                      df = degrees_freedom)

# generate a plot for each metabolite
for (metab in 1:length(metab_names)) {

  p <- santaR_plot(sp[[metab_names[metab]]],
                   showIndPoint=FALSE,
                   showIndCurve=FALSE,
                   xlab = NULL,
                   ylab = NULL,
                   colorVect = unlist(color_pal[-2]),
                   title = plot_titles[metab]) +
    scale_x_continuous(breaks = 1:length(tp_labs),
                       labels = tp_labs) +
    theme(plot.title = element_text(size = 6),
          legend.position = "None",
          axis.text = element_text(size = 5),
          panel.grid.major = element_blank(),
          panel.grid.minor = element_blank()) +
    # add p value
    annotate("text",
             size = 2,
             x = 5,
             y = pval_locs[metab],
             label = paste("P = ", round(sp[[metab_names[metab]]]$general$pval.dist, digits = 4)))

  assign(metab_names[metab], p)
}
```

(1b) Relative concentration of PAG was also higher in neonatal urine from malabsorptive group.

```r
pb <- ggboxplot(integrals,
                x = "Group_label",
                y = "PAG_B",
                color = "Group_label",
                ylim = c(-1,11),
                add = "jitter",
                size = 0.25,
                add.params = list(size = 0.5),
                palette = unlist(color_pal),
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
                legend = "none",
                xlab = FALSE,
                ylab = "Integral (a.u.)",
                title = "Neonatal Urine - PAG") +
  theme(axis.text=element_text(size = 5),
        axis.title = element_text(size = 6),
        plot.title = element_text(size = 8, hjust = 0.5)) +
  # add significance
  stat_compare_means(comparisons = list(c("NBS", "MAL"), c("RES", "MAL")),
                     label.y = c(7, 9.5),
                     label.y.npc = "bottom",
                     label =  "p.signif",
                     method = "wilcox.test")
```

Combine maternal (1a) serum, urine, and neonatal (1b) urine plots

```
p <- ggarrange(leucine, isoleucine, isobutyrate ,NAGP, glutamine, THB,
               PAG, PCS, IS, PHPA, unkn, valine, aKIV, MM, creatinine, pb)

fig1 <- annotate_figure(p,
                     left = text_grob("Integral (a.u.)",
                                      size = 10, rot = 90),
                     bottom = text_grob("Time point",
                                        size = 10))
```

Table S5: Maternal metabolic changes depend on type of bariatric surgery

Relative concentration changes for each discriminatory metabolite were detailed at all time points for both malabsorptive and restrictive groups.

```
################## DEFINED VARIABLES ###################

# input dataframe has one row for each sample, one column for each metabolite (with concentrations or integrals)
input_data <- integrals

# metabolites of interest (match column names in input data)
metab_names <- c("PAG", "PCS", "IS", "PHPA", "unkn", "creatinine", "aKIV", "MM", "valine",
                 "isoleucine", "leucine", "isobutyrate", "glutamine", "THB", "NAGP")

# column name for time point
tp_col <- "Time_point_label"

# time points to include
time_points <- c("T1", "T2", "T3", "T4", "T5", "T6")

# column name for group
group_col <- "Group"

# case group name
case_names <- c("Restrictive", "Malabsorptive") # function only takes one input group, use for loop

# control group name
con_name <- "No Bariatric Surgery"

############################################################
```

Generate table: log2(Fold Change) was calculated where Fold Change (FC) = mean(case)/mean(control). *P* values are from Mann-Whitney U test.

```
univariate_metabolites <- function(input_data,
                                   metab_names,
                                   tp_col,
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
                                    time_points,
                                    group_col,
                                    case_name,
                                    con_name) {
  # subset data to comparison groups
  input_data <- subset(input_data, input_data[[group_col]] %in% c(case_name, con_name))

  # create empty matrices to store results
  log2FC <- as.data.frame(matrix(nrow = length(time_points), ncol = length(metab_names))) %>%
    set_colnames(metab_names) %>%
    set_rownames(time_points)

  pval <- as.data.frame(matrix(nrow = length(time_points), ncol = length(metab_names))) %>%
    set_colnames(metab_names) %>%
    set_rownames(time_points)

  # Calculate log2(FC) for each metabolite at time points T1 - T6
  for (time in time_points){
    data <- subset(input_data, input_data[[tp_col]] == time)
    con <- subset(data, data[[group_col]] == con_name)
    case <- subset(data, data[[group_col]] == case_name)
    for (metab in metab_names){
      log2fc <- round(log2(mean(case[,metab], na.rm = T)/mean(con[,metab], na.rm = T)), digits = 3)
      log2FC[time,metab] <- log2fc
    }}

  # Calculate p value for each metabolite at time points T1 - T6
  for (time in time_points){
    for (metab in metab_names){
      data <- subset(input_data,
                     input_data[[tp_col]] == time & input_data[[group_col]] %in% c(case_name, con_name))
      form <- formula(paste(metab, "~", group_col))
      pval[time,metab] <- wilcox.test(form, data, exact = FALSE)$p.value
    }}

  # adjust p values
  log2FC.list <- data.frame(Time_point = rep(time_points, length(metab_names)),
                            bind_cols(gather(log2FC, "Metabolite", "log2FC"),
                                      gather(pval, "Metab", "p.val")))
  log2FC.list$p.adj.BH <- p.adjust(log2FC.list$p.val, method = "BH")
  log2FC.list$Group <- case_name
  return(log2FC.list[,-which(colnames(log2FC.list) == "Metab")])
}
```

Generate table for each case group

```
tableS5 <- NULL
for (gr in case_names) {
  idx <- univariate_metabolites(input_data,
                                metab_names,
                                tp_col,
                                time_points,
                                group_col,
                                case_name = gr,
                                con_name)
  tableS5 <- rbind(tableS5, idx)
}
```

Format table

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
# change p values to scientific notation
tableS5_formatted <- tableS5 %>%
  mutate(P = formatC(tableS5$p.val, format = "e", digits = 2),
            Padj = formatC(tableS5$p.adj.BH, format = "e", digits = 2))

# change non-significant p values to NS
tableS5_formatted$P[as.numeric(tableS5_formatted$P) > 0.05] <- "NS"
tableS5_formatted$Padj[as.numeric(tableS5_formatted$Padj) > 0.05] <- "NS"
```

Figure S3: Maternal metabolic changes depend on type of bariatric surgery

Graphical representation of the changes

```
################## DEFINED VARIABLES ##################

input_data <- tableS5_formatted %>%
  # label significant data
  mutate(plot_sig = ifelse(Padj == "NS", NA, "*"),
         Group_label = case_when(Group == "Malabsorptive" ~ "MAL",
                                 Group == "Restrictive" ~ "RES"),
         # Change names for plot titles
         Metabolite = plyr::mapvalues(Metabolite,
                                      from = c("THB", "NAGP", "unkn", "aKIV", "MM", "glutamine",
                                               "creatinine", "isobutyrate", "leucine", "isoleucine",
                                               "valine"),
                                      to = c("D-beta-hydroxybutyrate", "GlycA",
                                             "Unknown", "alpha-Ketoisovalerate", "Methylmalonate",
                                             "Glutamine", "Creatinine", "Isobutyrate", "Leucine",
                                             "Isoleucine", "Valine")),
         # order time points
         Time_point = factor(Time_point, levels = c("T6", "T5", "T4", "T3", "T2", "T1")),
         # add which biofluid metabolites were identified in
         biofluid = ifelse(Metabolite %in% c("PAG", "PCS", "IS", "PHPA", "Unknown", "Valine",
                                             "Creatinine", "alpha-Ketoisovalerate", "Methylmalonate"),
                           "urine",
                           "serum"))

# column names
tp_col <- "Time_point"
metab_col <- "Metabolite"
s_type_col <- "biofluid"
group_col <- "Group_label"
plot_val <- "log2FC"
label_sig <- "plot_sig"

figure_labels <- list(serum = "a", urine = "b")

###########################################################
```

Plot effect sizes and significance

```
for (s_type in names(figure_labels)) {

    # subset to appropriate biofluid type
    to_plot <- subset(input_data, input_data[[s_type_col]] == s_type)

    # plot log2FC for both groups
    p <- ggdotchart_not_ordered(to_plot,
                                x = tp_col,
                                xlab = "",
                                y = plot_val,
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
                                ylab = "",
                                rotate = TRUE,
                                color = group_col,
                                palette = unlist(color_pal),
                                dot.size = 1.5,
                                label = label_sig,
                                font.label = list(color = "black",
                                                  size = 10,
                                                  vjust = 0.72),
                                add = "segments",
                                add.param = list(color = "lightgray",
                                                 size = 0.4)) +
      ylim(c(min(to_plot$log2FC), max(to_plot$log2FC))) +
      geom_hline(yintercept = 0,
                 linetype = 2,
                 size = 0.4,
                 color = "lightgray") +
      theme(legend.title = element_blank(),
            legend.text=element_text(size = 6),
            legend.key.size = unit(3, "mm"),
            axis.text=element_text(size = 5),
            axis.line = element_line(size = 0.4),
            axis.ticks = element_line(size = 0.4)) +
      rotate_x_text(55)

    # separate plots for each metabolite
    p <- facet(p,
               facet.by = metab_col,
               nrow = 2,
               scales = "free_x",
               panel.labs.background = list(fill = "ivory2",
                                            color = "black"),
               panel.labs.font = list(size = 4,
                                      face ="bold"))

    # add figure labels and axis annotations
    p <- annotate_figure(p,
                         fig.lab.pos = "top.left",
                         fig.lab = figure_labels[[s_type]],
                         fig.lab.size = 10,
                         fig.lab.face = "bold",
                         left = text_grob("Time point",
                                          size = 6,
                                          rot = 90, vjust = 4),
                         bottom = text_grob("log2(Fold Change)",
                                            size = 6, vjust = -4))

    assign(paste0(s_type, "_dotplot"), p)
}
```

Combine serum and urine plots

```
figS3 <- ggarrange(serum_dotplot, urine_dotplot,
                  ncol = 2,
                  widths = c(1, 1.5))
```

Figure S7: Relative concentrations of urinary host-microbial co-metabolites vary depending on bariatric surgery sub-type

Comparison of key discriminatory metabolites between study sub-groups.

West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
################## DEFINED VARIABLES ##################

input_data <- integrals %>%
  # separate restrictive group into band and sleeve
  mutate(Group_label = case_when(Group_label == "NBS" ~ "NBS",
                                 Group_label == "MAL" ~ "MAL",
                                 Group_label == "RES" & Bariatric.surgery == "BAND" ~ "BAND",
                                 Group_label == "RES" & Bariatric.surgery == "SLEEVE" ~ "SLEEVE")) %>%
  # order sub-groups
  mutate(Group_label = factor(Group_label, levels = c("NBS", "BAND", "SLEEVE", "MAL"))) %>%
  # remove neonatal samples
  filter(Time_point_label != "T7")

# column names
metab_names <- c("PAG", "PCS", "IS", "PHPA", "unkn")
tp_col <- "Time_point_label"
group_col <- "Group_label"

colors <- list(NBS = "purple", BAND = "coral2", SLEEVE = "coral2", MAL = "darkturquoise")

###########################################################
```

For each metabolite, plot relative concentrations at each time point.

```
for (metab in metab_names) {
  for (tp in unique(input_data[[tp_col]])) {
    p <- ggboxplot(subset(input_data, input_data[[tp_col]] == tp),
                   size = 0.2,
                   group_col,
                   metab,
                   color = group_col,
                   palette = unlist(colors),
                   add = "jitter",
                   add.params = list(alpha = 0.5,
                                     size = 0.3),
                   xlab = FALSE,
                   ylab = paste(metab, tp)) +
      theme(legend.position = "None",
            axis.text=element_text(size = 5),
            axis.title = element_text(size = 6),
            axis.line = element_line(size = 0.4),
            axis.ticks = element_line(size = 0.4)) +
      rotate_x_text(angle = 45)
    p$layers[[1]]$geom_params$outlier.size <- 0.3 # change size of outlier points
    assign(paste0(metab, "_", tp, "_boxplot"), p)
  }
}
```

Combine all plots together

```
figS7 <- ggarrange(PAG_T1_boxplot, PAG_T2_boxplot, PAG_T3_boxplot, PAG_T4_boxplot,
                   PAG_T5_boxplot, PAG_T6_boxplot, PCS_T1_boxplot, PCS_T2_boxplot,
                   PCS_T3_boxplot, PCS_T4_boxplot, PCS_T5_boxplot, PCS_T6_boxplot,
                   IS_T1_boxplot, IS_T2_boxplot, IS_T3_boxplot, IS_T4_boxplot,
                   IS_T5_boxplot, IS_T6_boxplot, PHPA_T1_boxplot, PHPA_T2_boxplot,
                   PHPA_T3_boxplot, PHPA_T4_boxplot, PHPA_T5_boxplot, PHPA_T6_boxplot,
                   unkn_T1_boxplot, unkn_T2_boxplot, unkn_T3_boxplot, unkn_T4_boxplot,
                   unkn_T5_boxplot, unkn_T6_boxplot)
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

Figure 2: Gut bacterial alterations in malabsorptive patients

```
################## DEFINED VARIABLES ##################

# input data is phyloseq object
input_data <- ps_M

# column names
tp_col <- "Time_point_label"
group_col <- "Group_label"

adiv_measure <- "Shannon"
bdiv_measure <- "bray"
ordination_meth <- "MDS"

set.seed(71)

###########################################################
```

Calculate alpha-diversity

```
# normalize by rarefying to minimum sequencing depth
data <- input_data %>%
  rarefy_even_depth(sample.size = min(sample_sums(input_data)), verbose = FALSE, replace = FALSE)

# calculate alpha diversity
alpha_div <- estimate_richness(data, split = TRUE, measures = adiv_measure)

# add variables needed for plot
alpha_div$Time_point <- sample_data(data)[[tp_col]]
alpha_div$Group <- sample_data(data)[[group_col]]
```

Calculate beta-diversity

```
# use log transformation normalization to overcome differences in sequencing depth
data <- input_data %>%
  transform_sample_counts(function(x) {log(x+1)})

bdiv <- phyloseq::distance(data, bdiv_measure)
MDS <- ordinate(data, ordination_meth, bdiv)

# extract % variance explained for axis labels
ord <- plot_ordination(data, MDS)[["labels"]]
```

(2a) Comparison of Shannon Index between malabsorptive and control groups

```
am <- facet(ggboxplot(alpha_div,
               x = "Group",
               y = adiv_measure,
               color = "Group",
               add = "jitter",
               add.params = list(alpha = 0.5),
               size = 0.5,
               palette = unlist(color_pal),
               legend = "none",
               ylim = c(3.3,5.2),
               xlab = FALSE,
               ylab = "Shannon's Diversity Index (H)") +
              theme(axis.text = element_text(size = 6),
                    axis.title = element_text(size = 6)) +
              # add significance
              stat_compare_means(comparisons = list(c("NBS", "MAL")),
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
                                  label =  "p.signif",
                                  method = "wilcox.test",
                                  label.y = 5.1,
                                  hide.ns = TRUE),
            facet.by = "Time_point",
            panel.labs.background = list(fill = "ivory2", color = "black"),
            panel.labs.font = list(size = 7, face ="bold"))
```

(2b) Bray-Curtis dissimilarity with PCoA ordination

```
# plot PCoA
to_plot <- plot_ordination(data, MDS, justDF = TRUE)
bm <- ggscatter(to_plot,
                x = "Axis.1",
                y = "Axis.2",
                color = "Group_label",
                alpha = 0.6,
                ggtheme = theme_bw(),
                shape = "Time_point_label",
                size = 1.5,
                ellipse = TRUE,
                ellipse.level = 0.95,
                ellipse.type = "norm",
                ellipse.alpha = 0) +
  scale_colour_manual(values = unlist(color_pal)) +
  labs(x = ord$x,
       y = ord$y) +
  theme(legend.position = "none",
        axis.title =element_text(size=6),
        axis.text = element_text(size = 6))

# permanova test
adonis_res <- adonis(bdiv ~ Group, data.frame(sample_data(data)))

# add permanova p value
bm <- bm +
  annotate("text",
           size = 2,
           x = 0.3,
           y = 0.32,
           label = paste("P = ", round(adonis_res$aov.tab$`Pr(>F)`[1], digits = 4)))

# put 2a and 2b together
plot1 <- ggarrange(am, bm, labels = c("a", "b"), widths = c(2,1))
```

(2c) Differential abundance (malabsorptive vs control) at each taxonomic level. Effect sizes are log2(FC) and *P* values are from Mann-Whitney U test.

```
################## DEFINED VARIABLES ##################

# input data is phyloseq object
data <- ps_M

tp_col <- "Time_point_label"
time_points <- c("T1", "T2", "T4")

group_col <- "Group_label"
case_name <- "MAL"
con_name <- "NBS"

# 16S data transformation parameters
```

```
# taxa without at least min_counts in at least prev_filt (%) of samples will be removed
min_counts <- 5
prev_filt <- 0.1

############################################################
```

Generate differential abundance table (Table S6)

```
MWU <- NULL

for (rank in c("Phylum", "Class", "Order", "Family", "Genus")){
  for (time in time_points){

    ps_obj <- data %>%
      tax_glom(rank) %>%
      # prevalence filter
      filter_taxa(function(x) sum(x > min_counts) > (prev_filt*length(x)), TRUE) %>%
      # normalize by total sum scaling
      transform_sample_counts(function(x) x/sum(x)) %>%
      # subset to time point
      subset_samples(Time_point_label == time)

    tax <- data.frame(as(tax_table(ps_obj), "matrix"))
    abund <- data.frame(as(otu_table(ps_obj), "matrix")) %>%
      # change column names to names at current rank
      set_colnames(gsub("/", ".", tax[,which(colnames(tax)==rank)]))
    # add group column
    abund$Group <- sample_data(ps_obj)[[group_col]]

    # test differences between groups for every taxon
    for (i in c(1:(ncol(abund)-1))) {
      taxa <- colnames(abund)[1:(ncol(abund)-1)]
      form <- formula(paste(taxa[i], "~ Group"))
      mw <- wilcox.test(form, abund, exact = FALSE)
      idx <- tibble(Rank = rank,
                    Time_point = time,
                    Taxon = taxa[i],
                    Mean_rel_abund_case = mean(subset(abund, abund$Group == case_name)[,i]),
                    Mean_rel_abund_cont = mean(subset(abund, abund$Group == con_name)[,i]),
                    pval = mw$p.value,
                    log2FC = log2(mean(subset(abund, abund$Group == case_name)[,i])/
                                    mean(subset(abund, abund$Group == con_name)[,i]))
      )
      MWU <- rbind(MWU, idx)
    }
  }
}

# multiple hypothesis testing correction
MWU$p.adj.BH <- p.adjust(MWU$pval, method = "BH")

# filter to significant taxa; excluding taxa that are not observed in a group (zero counts)
MWU_filt <- subset(MWU, MWU$p.adj.BH < 0.05 & abs(MWU$log2FC) != Inf)
```

Barplots for each rank

```
# color related taxa
palette <- list(Proteobacteria = "cyan4",
                Bacilli = "khaki4",
                Actinomycetales = "chocolate3",
                Anaerostipes = "thistle4",
```

```
                NS = "grey")

diff_abund <- MWU %>%
  # add x-axis plot label
  mutate(Label = str_replace(paste0(Taxon, "_", Time_point),
                            "Escherichia\\.", "Escherichia/"),
         # group taxa for coloring
         Taxa_group = case_when(p.adj.BH > 0.05 ~ "NS",
                               grepl("Proteobact|Gammaprot|Enterobacteria|Escherichia|Pasteurell|Haemophil",
                                     Taxon) ~ "Proteobacteria",
                               Taxon %in% c("Micrococcaceae", "Rothia", "Actinomycetales") ~ "Actinomycetales",
                               grepl("Bacill|Enteroc|Strepto|Lactobacill|Carnobact|Granulicat",
                                     Taxon) ~ "Bacilli",
                               grepl("Bacteroid|Butyricimonas", Taxon) ~ "Bacteroidetes",
                               Taxon == "Anaerostipes" ~ "Anaerostipes",
                               grepl("Holdeman|Solobact", Taxon) ~ "Erysipelotrichaceae",
                               grepl("Veillonella|Acidaminococcus", Taxon) ~ "Negativicutes",
                               grepl("Clostrid", Taxon) ~ "Clostridia"))

# factor for correct ordering on plot
diff_abund$Rank <- factor(diff_abund$Rank, levels = c("Phylum", "Class", "Order", "Family", "Genus"))
diff_abund$Taxa_group <- factor(diff_abund$Taxa_group, levels = c("Proteobacteria",
                                                                  "Bacilli",
                                                                  "Bacteroidetes",
                                                                  "Clostridia",
                                                                  "Negativicutes",
                                                                  "Actinomycetales",
                                                                  "Erysipelotrichaceae",
                                                                  "Anaerostipes",
                                                                  "NS"))

# remove taxa only significant at one time point
MWU_plot <- subset(diff_abund, diff_abund$Taxon %in%
                              names(which(table(MWU_filt$Taxon) > 1)))

# top barplots
p <- ggbarplot(subset(MWU_plot, MWU_plot$Rank %in% c("Phylum", "Class", "Order")),
              x = "Label",
              y = "log2FC",
              lab.size = 4,
              fill = "Taxa_group",
              color = "white",
              palette = unlist(palette),
              ylab = "log2(Fold Change)",
              lab.vjust = .5,
              xlab = FALSE,
              order = sort(MWU_plot$Label)) +
  theme(axis.title = element_text(size=6),
        axis.text = element_text(size = 6),
        legend.position = "none") +
  rotate_x_text(55)

p <- facet(p,
           facet.by = "Rank",
           scales = "free_x",
           nrow = 1,
           space = "free_x",
           panel.labs.background = list(fill = "ivory2", color = "black"),
           panel.labs.font = list(size = 8, face ="bold"))
```

```
# bottom barplots
q <- ggbarplot(subset(MWU_plot, MWU_plot$Rank %in% c("Family", "Genus")),
              x = "Label",
              y = "log2FC",
              lab.size = 4,
              fill = "Taxa_group",
              color = "white",
              palette = unlist(palette),
              ylab = "log2(Fold Change)",
              lab.vjust = .5,
              xlab = FALSE,
              order = sort(MWU_plot$Label)) +
  theme(axis.title = element_text(size=6),
        axis.text = element_text(size = 6),
        legend.position = "right",
        legend.text = element_text(size = 6),
        legend.title = element_text(size = 8),
        legend.key.size = unit(4, "mm")) +
  rotate_x_text(55) +
  geom_hline(yintercept = 0, linetype="dashed", size = 0.4)

q <- facet(q,
           facet.by = "Rank",
           scales = "free_x",
           nrow = 1,
           space = "free_x",
           panel.labs.background = list(fill = "ivory2", color = "black"),
           panel.labs.font = list(size = 7, face ="bold"))

# combine barplots
plot2 <- ggarrange(p, q, nrow = 2)
```

Combine (2c) with (2a) and (2b)

```
fig2 <- ggarrange(plot1, plot2,
                  nrow = 2,
                  heights = c(1.2,2),
                  labels = c("", "c"))
```

Table S7: Differentially abundant amplicon sequence variants

Differential abundance of amplicon sequence variants was assessed by DESeq2. Taxonomic assignments at species level were compared between SILVA and RDP databases.

```
################## DEFINED VARIABLES ##################

input_data <- merge_samples(ps_M, "Study.no")

# fix group labels - merging samples changes labels to numbers
sample_data(input_data)$Group <- str_replace_all(sample_data(input_data)$Group,
                                                 c("1" = "Control",
                                                 "2" = "Malabsorptive"))
sample_data(input_data)$Group <- factor(sample_data(input_data)$Group, levels = c("Control", "Malabsorptive"))

# 16S data transformation parameters
# taxa without at least min_counts in at least prev_filt (%) of samples will be removed
min_counts <- 5
prev_filt <- 0.1

alpha <- 0.05
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
gen <- c("Enterococcus", "Escherichia/Shigella", "Streptococcus", "Rothia")

###########################################################
```

Make table with differential abundance and species assignments

```
# prevalence filter
ps_trim <- input_data %>%
  filter_taxa(function(x) sum(x > min_counts)> (prev_filt*length(x)), TRUE)

# run DESeq2
DDS <- phyloseq_to_deseq2(ps_trim, ~ Group)
DDS <- estimateSizeFactors(DDS, type = "poscounts")
DDS <- estimateDispersions(DDS, fitType = "local")
DDS <- DESeq(DDS, fitType = "local")

# results
res <- results(DDS)
res <- res[order(res$padj, na.last=NA), ]

sigtab <- res[(res$padj < alpha), ]
sigtab <- cbind(as(sigtab, "data.frame"), as(tax_table(ps_trim)[rownames(sigtab), ], "matrix"))

# keep only genera of interest
sigtab <- subset(sigtab, sigtab$Genus %in% gen)
keep <- which(row.names(rdp.species) %in% row.names(sigtab))

rdp.species <- data.frame(rdp.species)
rdp.species$seq <- row.names(rdp.species)
silva.species <- data.frame(silva.species)
silva.species$seq <- row.names(silva.species)

# combine RDP and SILVA species assignments
species_assign <- merge(rdp.species[keep, c(6:8)], silva.species[keep, c(6:8)], by = "seq")

# combine DESeq2 results with species assignments
sigtab$seq <- row.names(sigtab)
species_assign <- merge(sigtab[,c(1:6,13)], species_assign, by = "seq")
```

Figure 3: Integration of serum, urine and faecal profiles

Relative concentrations of key discriminatory metabolites were integrated with genera abundances using "DIABLO", a method implemented in `mixOmics`.

```
################## DEFINED VARIABLES ##################

# input data is phyloseq object
input_MB <- ps_M %>%
  tax_glom("Genus") %>%
  transform_sample_counts(function(x) {log(x+1)})

# remove neonatal samples & subset to individuals with microbiome data
input_metab <- integrals %>%
  filter(Time_point != 7)
input_metab <- left_join(data.frame(Patient.TP = sample_names(input_MB)),
                         input_metab,
                         by = "Patient.TP")

S_metab_names <- c("leucine", "isoleucine", "isobutyrate", "THB")
U_metab_names <- c("PAG", "PCS", "IS", "PHPA", "unkn", "aKIV", "creatinine")
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
# number of genera to keep in model
n_gen <- 10

############################################################
```

Prepare datasets

```
tax <- as(tax_table(input_MB), "matrix") %>%
  data.frame(stringsAsFactors = F)

F_abund_mat <- as(otu_table(input_MB), "matrix") %>%
  data.frame() %>%
  set_colnames(tax$Genus)

S_NMR_mat <- input_metab %>%
  select(one_of(S_metab_names)) %>%
  set_rownames(input_metab$Patient.TP)

U_NMR_mat <- input_metab %>%
  select(one_of(U_metab_names)) %>%
  set_rownames(input_metab$Patient.TP)
```

Build model

```
X <- list(Urine = U_NMR_mat,
          Serum = S_NMR_mat,
          Faeces = F_abund_mat)
Y <- input_metab[, "Group_label"]

# only model top n taxa for each component
list.keepX <- list(Urine = ncol(U_NMR_mat), Serum = ncol(S_NMR_mat), Faeces = c(n_gen, n_gen))

# multi-block pls-da
m <- block.splsda(X, Y,
                  keepX = list.keepX,
                  near.zero.var = TRUE)
```

Plot sample scores for each dataset. Separation between classes is on Component 1.

```
ggsave(plotIndiv(m,
                 legend = TRUE,
                 X.label = "Component 1",
                 Y.label = "Component 2",
                 legend.title = '',
                 style = "lattice",
                 ind.names = FALSE,
                 col.per.group = unlist(color_pal)[-2],
                 ellipse = TRUE,
                 legend.position = "top",
                 pch = 20, alpha =  0.7,
                 size.axis = 0.5, size.xlabel = 0.9, size.ylabel = 0.9,
                 size.legend = 0.8) +
         theme(legend.box.background = element_rect()),
       filename = file.path(save_dir, "figure3a_integration_scores.pdf"),
       height = 3, width = 6)
```

Plot correlations between variables and components highlighting variables with high correlation on either component.

```
var_plot <- plotVar(m, plot = F) %>%
  mutate(Plot_label = plyr::mapvalues(names,
                                      from = c("THB", "unkn", "aKIV", "creatinine",
                                               "isobutyrate", "leucine", "isoleucine"),
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
                                          to = c("D-beta-HB", "Unknown",
                                                 "alpha-Ketoisovalerate", "Creatinine",
                                                 "Isobutyrate", "Leucine", "Isoleucine")),
         Plot_label = case_when(abs(x) > 0.5 | abs(y) >= 0.5 ~ Plot_label),
         Plot_alpha = ifelse(is.na(Plot_label), "|r|<0.5", "|r|>=0.5"))

ggsave(ggscatter(var_plot, x = "x", y = "y",
                 color = "Block", palette = c("darkolivegreen", "darkgoldenrod3", "rosybrown3"),
                 label = "Plot_label", repel = TRUE,
                 font.label = c(5, "plain"),
                 label.rectangle = FALSE,
                 alpha = "Plot_alpha",
                 xlab = "Component 1", ylab = "Component 2",
                 show.legend.text = FALSE) +
         xlim(c(-1, 1)) +
         ylim(c(-1,1)) +
         geom_hline(yintercept = 0,
                    linetype = 2,
                    size = 0.4,
                    color = "lightgray") +
         geom_vline(xintercept = 0,
                    linetype = 2,
                    size = 0.4,
                    color = "lightgray") +
         theme(legend.title = element_blank(),
               axis.text = element_text(size = 5),
               axis.title = element_text(size = 8),
               legend.text = element_text(size = 5)),
       filename = file.path(save_dir, "figure3b_integration_variables.pdf"),
       height = 3.5, width = 3.5)
```

Plot correlations between datasets for Component 1.

```
ggsave(circosPlot(m,
                  cutoff=0.5,
                  comp = 1,
                  color.cor = c("red", "blue"),
                  color.blocks = c("rosybrown3", "darkgoldenrod3", "darkolivegreen"),
                  showIntraLinks = FALSE,
                  legend = TRUE,
                  line = FALSE,
                  size.labels = 0.01,
                  size.variables = 0.5,
                  var.names = list(Urine = c("PAG","PCS","IS","PHPA","Unknown","alpha-KIV","Creatinine"),
                                   Serum = c("Leucine", "Isoleucine", "Isobutyrate", "D-beta-HB"),
                                   Faeces = m$names$colnames$Faeces),
                  size.legend = 0.5) +
         theme(legend.box.background = element_rect()),
       filename = file.path(save_dir, "figure3c_integration_circos.pdf"),
       height = 4, width = 4.5)
```

Figure 4: Relative concentrations of metabolites correlate with maternal insulin resistance and fetal/birth weight

Correlations (Spearman) were calculated between metabolite relative concentrations and clinical measures. Netowrk file was generated for Cytoscape input.

```
################## DEFINED VARIABLES ##################

time_points <- 1:6
input_data <- subset(integrals, integrals$Time_point %in% time_points)
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
tp_col <- "Time_point_label"

# column names
metab_names <- c("PAG", "PCS", "IS", "isoleucine", "leucine", "isobutyrate")
clin_varT1 <- c("Fast Maternal Insulin (microU/ml)",
                "Maternal HOMA-IR",
                "Fast Maternal Gluc (mmol/L)",
                "GA..days.",
                "BMI.1",
                "Age")
clin_varT2 <- c("Fast Maternal Insulin (microU/ml)",
                "Maternal HOMA-IR",
                "Fast Maternal Gluc (mmol/L)",
                "GA..days.",
                "BMI2",
                "EFW2-centile",
                "Age")
clin_varT3 <- c("Fast Maternal Insulin (microU/ml)",
                "Maternal HOMA-IR",
                "Fast Maternal Gluc (mmol/L)",
                "GA..days.",
                "BMI3",
                "Age")
clin_varT4 <- c("Fast Maternal Insulin (microU/ml)",
                "Maternal HOMA-IR",
                "Fast Maternal Gluc (mmol/L)",
                "GA..days.",
                "BMI4",
                "EFW4-centile",
                "Age")
clin_varT5 <- c("Fast Maternal Insulin (microU/ml)",
                "Maternal HOMA-IR",
                "Fast Maternal Gluc (mmol/L)",
                "GA..days.",
                "BMI5",
                "EFW5-centile",
                "Age")
clin_varT6 <- c("Fast Maternal Insulin (microU/ml)",
                "Maternal HOMA-IR",
                "Fast Maternal Gluc (mmol/L)",
                "GA..days.",
                "BMI6",
                "BW.centile",
                "Age")

############################################################
```

Generate network file

```
# use T5 BMI for T6 (visits are couple weeks apart, BMI not measured at T6)
input_data$BMI6 <- input_data$BMI5

# format data for correlation at each time point
for (time in unique(input_data[[tp_col]])) {
  idx <- subset(input_data, input_data[[tp_col]] == time)
  # keep only columns to correlate
  idx <- idx[,which(colnames(idx) %in% c(metab_names, get(grep(paste0("clin_var", time), ls(), value = T))))]
  # fix column names so they all match
  colnames(idx) <- str_replace_all(colnames(idx), c("BMI.*" = "BMI", ".*centile" = "EFW.BW"))
  if (!"EFW.BW" %in% colnames(idx)) { idx$EFW.BW <- NA }
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
  # tidy column names
  colnames(idx) <- make.names(colnames(idx))
  # arrange columns in the same order
  idx <- idx[,c(metab_names,
                "Age",
                "BMI",
                "GA..days.",
                "Fast.Maternal.Insulin..microU.ml.",
                "Fast.Maternal.Gluc..mmol.L.",
                "Maternal.HOMA.IR",
                "EFW.BW")]
  # make sure all columns are numeric
  idx <- sapply(idx, as.numeric)
  assign(paste(time), idx)
}

# remove confounders
for (time in unique(input_data[[tp_col]])) {
  idx <- get(paste(time))
  assign(paste0(time, "_to_corr"), idx[,-c(which(colnames(idx) %in% c("Age", "BMI")))])
}

# correlation at each time point
corr_all <- NULL
corr_pval <- NULL
for (cor in time_points) {
  idx <- get(grep(paste0(cor,"_to_corr"), ls(), value = T))
  rcorr <- rcorr(idx, type = "spearman")
  coef <- round(rcorr$r,2)
  coef <- coef[c(which(rownames(coef) %in% metab_names)), c(which(!colnames(coef) %in% metab_names))]
  row.names(coef) <- paste0(metab_names, "_T", cor)
  pval <- rcorr$P
  pval <- pval[c(which(rownames(pval) %in% metab_names)), c(which(!colnames(pval) %in% metab_names))]
  row.names(pval) <- paste0(metab_names, "_T", cor)
  corr_all <- rbind(corr_all, coef)
  corr_pval <- rbind(corr_pval, pval)
}

# change to long format and adjust p values
corr_all <- melt(corr_all, value.name = "spearman")
corr_all <- subset(corr_all, !is.na(corr_all$spearman))
corr_pval <- melt(corr_pval, value.name = "p.val")
corr_pval <- subset(corr_pval, !is.na(corr_pval$p.val))

corr_pval$p.adj.BH <- p.adjust(corr_pval$p.val, method = "BH")
corr_metaVmetab <- cbind(corr_all, corr_pval[,c(3:4)])
corr_metaVmetab_sig <- subset(corr_metaVmetab, corr_metaVmetab$p.adj.BH < 0.05)

# convert to graph for cytoscape visualisation
cytoscape <- corr_metaVmetab_sig[,c(1:3)] %>%
  mutate(Strength = case_when(abs(spearman) < 0.4 ~ "0.3-0.4",
                              abs(spearman) >= 0.4 ~ "0.4-0.5"),
         Direction = case_when(spearman > 0 ~ "pos",
                               spearman < 0 ~ "neg")) %>%
  rename(from = Var1, to = Var2, weight = spearman)

net <- graph.data.frame(cytoscape, directed = FALSE)
write_graph(net, file.path(save_dir, "cormat_cytoscape.gml"), format = "gml")
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

Table S8: Relative concentrations of metabolites correlate with maternal insulin resistance and fetal/birth weight

Includes partial correlations to adjust for the effect of maternal age and BMI as well as any effect of maternal insulin resistance on the baby's weight.

```
corr_metaVmetab_sig$partial <- NA
corr_metaVmetab_sig$part.pval <- NA

for (part_cor in c(1:nrow(corr_metaVmetab_sig))) {
  var1 <- gsub("_T.*", "", corr_metaVmetab_sig[part_cor,1])
  var2 <- corr_metaVmetab_sig[part_cor,2]
  idx <- get(paste0(gsub(".*_", "", corr_metaVmetab_sig[part_cor,1])))

  if (var2 == "EFW.BW") {
    conf <- which(colnames(idx) %in% c("Age", "BMI", "Maternal.HOMA.IR"))
    # remove any rows with missing data
    idx <- subset(idx,
                  !is.na(idx[,which(colnames(idx)==var1)]) &
                  !is.na(idx[,which(colnames(idx)==var2)]) &
                  !is.na(idx[,which(colnames(idx)=="Maternal.HOMA.IR")]) &
                  !is.na(idx[,which(colnames(idx)=="BMI")]) &
                  !is.na(idx[,which(colnames(idx)=="Age")]))
  } else {
    conf <- which(colnames(idx) %in% c("Age", "BMI"))
    # remove any rows with missing data
    idx <- subset(idx,
                  !is.na(idx[,which(colnames(idx)==var1)]) &
                  !is.na(idx[,which(colnames(idx)==var2)]) &
                  !is.na(idx[,which(colnames(idx)=="BMI")]) &
                  !is.na(idx[,which(colnames(idx)=="Age")]))
  }

  # partial correlations
  part_corr <- pcor.test(idx[,which(colnames(idx)==var1)],
                         idx[,which(colnames(idx)==var2)],
                         idx[,conf],
                         method = "spearman")
  corr_metaVmetab_sig$partial[part_cor] <- part_corr$estimate
  corr_metaVmetab_sig$part.pval[part_cor] <- part_corr$p.value

  corr_metaVmetab_sig$part.pval.adj <- p.adjust(corr_metaVmetab_sig$part.pval, method = "BH")
}

for (col in c("p.val","p.adj.BH","part.pval","part.pval.adj")) {
  corr_metaVmetab_sig[as.numeric(corr_metaVmetab_sig[[col]]) > 0.05, col] <- "NS"
}
```

Figure S8: Key metabolites are not associated with percentage of weight lost or length of time since surgery in malabsorptive patients

```
to_cor <- integrals %>%
  filter(Group_label == "MAL",
         Time_point != 7)

for (variab in c("perc_wl", "Months_betw_op_concep")) {
  for (metab in c("PAG", "PCS", "IS")) {
    p <- ggscatter(to_cor,
                   x = variab,
                   xlab = str_replace_all(variab,
                                          c("perc_wl" = "Post-surgery weight loss (%)",
```

```
                                              "Months_betw_op_concep" =
                                                "Time between operation and conception (months)")),
                         y = metab,
                         ylab = paste(metab, "(a.u.)"),
                         facet.by = "Time_point_label",
                         alpha = 0.7,
                         size = 1,
                         color = "darkturquoise",
                         cor.coef = TRUE,
                         cor.coeff.args = list(method = "spearman"),
                         cor.coef.size = 1.5) +
      theme(axis.text = element_text(size = 6),
            axis.title = element_text(size = 7))
    assign(paste0(metab, "_", variab), p)
  }
}

p <- ggarrange(PAG_perc_wl, PCS_perc_wl, IS_perc_wl,
               PAG_Months_betw_op_concep, PCS_Months_betw_op_concep, IS_Months_betw_op_concep)
```

Session Info

```
sessionInfo()
```

```
## R version 3.5.0 (2018-04-23)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS  10.15.1
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_GB.UTF-8/en_GB.UTF-8/en_GB.UTF-8/C/en_GB.UTF-8/en_GB.UTF-8
##
## attached base packages:
## [1] parallel  stats4    stats     graphics  grDevices utils     datasets
## [8] methods   base
##
## other attached packages:
##  [1] plotly_4.8.0                reshape2_1.4.3
##  [3] ppcor_1.1                   igraph_1.2.4
##  [5] Hmisc_4.2-0                 Formula_1.2-3
##  [7] survival_2.44-1.1           mixOmics_6.6.2
##  [9] MASS_7.3-51.4               DESeq2_1.22.2
## [11] SummarizedExperiment_1.12.0 DelayedArray_0.8.0
## [13] BiocParallel_1.16.6         matrixStats_0.54.0
## [15] Biobase_2.42.0              GenomicRanges_1.34.0
## [17] GenomeInfoDb_1.18.2         IRanges_2.16.0
## [19] S4Vectors_0.20.1            BiocGenerics_0.28.0
## [21] vegan_2.5-4                 lattice_0.20-38
## [23] permute_0.9-5               ggpubr_0.2
## [25] magrittr_1.5                phyloseq_1.26.1
## [27] santaR_1.0                  forcats_0.4.0
## [29] stringr_1.4.0               dplyr_0.8.0.1
## [31] purrr_0.3.2                 readr_1.3.1
## [33] tidyr_0.8.3                 tibble_2.1.1
## [35] ggplot2_3.2.1               tidyverse_1.2.1
##
## loaded via a namespace (and not attached):
##  [1] colorspace_1.4-1       htmlTable_1.13.1       corpcor_1.6.9
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620

```
##  [4] XVector_0.22.0         base64enc_0.1-3        rstudioapi_0.10
##  [7] bit64_0.9-7            RSpectra_0.14-0        AnnotationDbi_1.44.0
## [10] lubridate_1.7.4        xml2_1.2.0             codetools_0.2-16
## [13] splines_3.5.0          doParallel_1.0.14      geneplotter_1.60.0
## [16] knitr_1.22             shinythemes_1.1.2      ade4_1.7-13
## [19] jsonlite_1.6           annotate_1.60.1        broom_0.5.2
## [22] cluster_2.0.8          shiny_1.3.0            compiler_3.5.0
## [25] httr_1.4.0             backports_1.1.3        assertthat_0.2.1
## [28] Matrix_1.2-17          lazyeval_0.2.2         cli_1.1.0
## [31] later_0.8.0            acepack_1.4.1          htmltools_0.3.6
## [34] tools_3.5.0            gtable_0.3.0           glue_1.3.1
## [37] GenomeInfoDbData_1.2.0 Rcpp_1.0.1             cellranger_1.1.0
## [40] Biostrings_2.50.2      multtest_2.38.0        ape_5.3
## [43] nlme_3.1-138           iterators_1.0.10       xfun_0.6
## [46] rvest_0.3.2            mime_0.6               XML_3.98-1.19
## [49] zlibbioc_1.28.0        scales_1.0.0           hms_0.4.2
## [52] promises_1.0.1         biomformat_1.10.1      rhdf5_2.26.2
## [55] RColorBrewer_1.1-2     yaml_2.2.0             memoise_1.1.0
## [58] gridExtra_2.3          rpart_4.1-13           RSQLite_2.1.1
## [61] latticeExtra_0.6-28    stringi_1.4.3          genefilter_1.64.0
## [64] foreach_1.4.4          checkmate_1.9.1        rlang_0.3.4
## [67] pkgconfig_2.0.2        bitops_1.0-6           evaluate_0.13
## [70] Rhdf5lib_1.4.3         htmlwidgets_1.3        bit_1.1-14
## [73] tidyselect_0.2.5       plyr_1.8.4             R6_2.4.0
## [76] generics_0.0.2         DBI_1.0.0              pillar_1.3.1
## [79] haven_2.1.0            foreign_0.8-71         withr_2.1.2
## [82] mgcv_1.8-28            RCurl_1.95-4.12        nnet_7.3-12
## [85] modelr_0.1.4           crayon_1.3.4           rARPACK_0.11-0
## [88] ellipse_0.4.1          rmarkdown_1.12         locfit_1.5-9.1
## [91] grid_3.5.0             readxl_1.3.1           data.table_1.12.2
## [94] blob_1.1.1             digest_0.6.18          xtable_1.8-3
## [97] httpuv_1.5.1           munsell_0.5.0          viridisLite_0.3.0
```


West KA, *et al. Gut* 2020;0:1–8. doi: 10.1136/gutjnl-2019-319620