



An explainable machine-learning model using baseline gut microbiota, circulating short-chain fatty acids, and clinical features to predict weight-loss response after Roux-en-Y gastric bypass

Juan José Hernández-Morante, PhD^{a,b,*}, María Antonia Martínez-Sánchez, PhD^{b,c}, Inmaculada Ros-Madrid, MD, PhD^{b,c}, Virginia Esperanza Fernández-Ruiz, PhD^{b,c}, Mercedes Ferrer-Gómez, MD, PhD^{b,c}, Antonio Silva, PhD^d, Andrés Balaguer-Román, MD, PhD^{b,e}, María Dolores Frutos, MD, PhD^e, Horacio Pérez-Sánchez, PhD^f, Bruno Ramos-Molina, PhD^b

Background: Roux-en-Y gastric bypass (RYGB) is one of the most effective treatments for severe obesity; however, post-operative weight-loss response remains heterogeneous. To address this variability, we developed an explainable machine-learning (ML) model integrating baseline gut microbiota composition, circulating short-chain fatty acids (SCFAs), and clinical features to predict weight-loss outcomes after RYGB.

Methods: In this prospective cohort study, 143 patients undergoing RYGB were followed for 12 and 24 months. The participants were classified as high- or low-responders based on achieving $\geq 30\%$ total weight loss (%TWL). Gut microbiota composition was assessed using 16S rRNA gene sequencing. Circulating SCFAs were quantified by liquid chromatography–mass spectrometry. Microbiome-only, clinical–metabolic, and combined models were developed using LASSO logistic regression. Internal validation was performed using stratified repeated five-fold cross-validation with 20 repetitions, and performance was estimated from participant-level out-of-fold predictions. Shapley Additive Explanations (SHAP) were used descriptively to interpret model behavior.

Results: At 12 and 24 months after RYGB, 94 and 83 participants achieved $\geq 30\%$ TWL, respectively. Baseline anthropometric parameters did not differ between the response groups. The baseline alpha diversity was higher in high responders at both 12 and 24 months ($P = 0.009$ and $P = 0.021$, respectively). At the taxonomic level, baseline *Beijerinckiaceae* and *Lactobacillaceae* were associated with weight loss outcomes. Clinical–metabolic and combined models showed good internal discrimination, particularly at 12 months. However, adding microbiome features did not produce a consistent improvement over the clinical–metabolic model. SHAP values highlighted inflammation-related markers (cholinesterase and C-reactive protein) as the main predictors of 12-month weight loss; heptanoate levels and *Coriobacteriaceae* abundance were also relevant contributors.

Conclusions: Baseline combined ML models integrating baseline gut microbiota, circulating SCFAs, and clinical features provided moderate, internally validated discrimination for marked weight-loss response after RYGB, with the strongest performance observed at 12 months. These findings support further evaluation of preoperative ML models for patient stratification. Microbiome and circulating SCFA features may provide complementary biological and interpretive information, but their incremental predictive and clinical value requires further validation.

Keywords: bariatric surgery, gut microbiota, machine learning, severe obesity, short-chain fatty acids

^aEating Disorders Research Unit, UCAM Catholic University of Murcia, Murcia, Spain, ^bObesity, Diabetes and Metabolism Research Group, Biomedical Research Institute of Murcia (IMIB), Murcia, Spain, ^cDepartment of Endocrinology and Nutrition, Virgen de la Arrixaca University Clinical Hospital, Murcia, Spain, ^dFacility of Innovation and Analysis in Animal Source Foodstuffs of SAIUEX, University of Extremadura, Cáceres, Spain, ^eDepartment of General and Digestive System Surgery, Virgen de la Arrixaca University Clinical Hospital, Murcia, Spain and ^fStructural Bioinformatics and High-Performance Computing Research Group (BIO-HPC), Computer Engineering Department, UCAM Catholic University of Murcia, Murcia, Spain

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

*Corresponding author. Address: Eating Disorders Research Unit, UCAM Catholic University of Murcia, Campus de los Jerónimos, s/n., 30107 Guadalupe, Murcia, Spain. Obesity, Diabetes and Metabolism Research Group, Biomedical Research Institute of Murcia (IMIB), Murcia 30107, Spain. Tel.: +34 968 278 808. E-mail: jjhernandez@ucam.edu (J.J. Hernández-Morante).

Copyright © 2026 The Author(s). Published by Wolters Kluwer Health, LLC. This is an open access article distributed under the terms of the Creative Commons

Introduction

Roux-en-Y gastric bypass (RYGB) is a bariatric surgery (BS) technique that is currently the best treatment option for severe obesity. Maximum weight loss is usually achieved 18 months after surgery, with an average of 33–35% of total body weight

Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

International Journal of Surgery (2026) 00:1–14

Received 16 February 2026; Accepted 19 August 2026

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.ijournal.com/international-journal-of-surgery.

Published online 17 September 2026

<http://dx.doi.org/10.1097/JS9.0000000000005650>

lost at 3 years^[1], and 10-year results show that most patients maintain an average loss of 28–30% of total body weight, which means that the RYGB effect is highly durable^[2]. Moreover, recent comparative studies indicate that RYGB is superior to sleeve gastrectomy in terms of weight loss and quality of life in the medium- and long-term^[3].

The RYGB technique not only promotes significant weight loss but also induces improvement or remission in type 2 diabetes and provides sustained renal benefits^[4]. Data from previous studies have shown complete remission rates of type 2 diabetes between 60 and 70%, with significant improvements in glycosylated hemoglobin (HbA1c) and blood glucose levels at both 1 and 5 years after surgery^[5]. In addition, long-term studies and systematic reviews have shown that RYGB results in a substantial and sustained reduction in liver fat, inflammation, and fibrosis^[6]. Resolution of steatosis occurs in up to 95% of cases after RYGB, with marked improvements in metabolic-associated steatohepatitis (MASH) and a significant decrease in fibrosis (26% reduction in severity, 53% reduction in steatohepatitis activity)^[6]. Clinical trials have shown that the probability of resolution of MASH without fibrosis progression is over four times higher following RYGB than with lifestyle modification alone^[7].

Overall, it seems clear that RYGB-induced weight loss is effective not only for the treatment of obesity but also for its associated comorbidities. Unfortunately, there is great variation in the amount of weight loss and metabolic improvement among patients undergoing this surgery. A multicenter prospective study of more than 2300 patients found an individual variance of 20% in the amount of weight lost after RYGB, which was associated with factors such as age, initial BMI, presence of type 2 diabetes, psychiatric disorders, and other comorbidities^[8]. A recent open-label randomized trial showed that the range of total body weight loss after RYGB varies significantly within the sample, depending on the clinical profile of each patient (e.g., success defined as $\geq 60\%$ excess weight loss is achieved in 71–74% of cases within 1–3 years, while the remainder experience suboptimal results)^[9]. Several genetic markers and other clinical traits have also been shown to predict a poorer response to weight loss, supporting the existence of interindividual variation^[10]. It is important, however, to distinguish variables measured before surgery that may have predictive value from biological changes observed after surgery, as well as from cross-sectional associations that do not necessarily imply prediction or causality.

The gut microbiota has been investigated from two distinct perspectives: as a baseline biological profile potentially associated with subsequent weight-loss response, and as a biological system that changes substantially after RYGB. Longitudinal studies have reported associations between greater microbial diversity and more favorable long-term weight maintenance outcomes^[11]; however, such associations do not necessarily establish independent predictive value. Profound and sustained changes in the microbiota have been observed after RYGB and are characterized by a reduction in Firmicutes and an increase in Bacteroidetes and Proteobacteria, and these shifts have been associated with a more favorable metabolic phenotype and better weight-loss results^[12]. However, other studies have reported an inverse relationship between these phyla^[13]. Remarkably, these postoperative changes represent biological responses to surgery and should be distinguished from the potential predictive value of microbiota profiles measured before surgery. In addition to changes at the phylum

HIGHLIGHTS

- Baseline gut microbiota composition and diversity differ between patients with high and low weight-loss responses after bariatric surgery, suggesting a potential role in surgical outcomes.
- Combining microbiota profiles at the family level with clinical and metabolic variables improves postoperative outcome modeling, particularly at 12 months.
- Machine-learning approaches uncover complex microbial and clinical relationships that are not readily captured by conventional analyses, while explainable models (Shapley Additive Explanations) enhance clinical interpretability.
- Microbiota-informed models may support patient stratification and personalized postoperative management in bariatric surgery, although external validation is required before clinical implementation.

level, specific microbial taxa have been associated with the effectiveness of RYGB. *Sarcina*, *Butyrivibrio*, *Alkaliphilus*, *Lachnospira*, *Pseudoalteromonas*, and *Cetobacterium* have been reported to be more abundant in individuals who experience greater weight loss following RYGB^[14]. Moreover, other studies have highlighted that *Akkermansia muciniphila*, as well as short-chain fatty acid (SCFA)-producing genera, such as *Lachnospiraceae* and *Dialister*, have been associated with more favorable postoperative metabolic outcomes, improved weight-loss outcomes^[15], and MASH resolution^[16].

Remarkably, not only microbiota composition but also the relationships between species appear to influence postsurgical outcomes. Specifically, a prospective study found that a higher proportion of *Prevotella* relative to *Bacteroides* (*Prevotella*-to-*Bacteroides* ratio) was associated with greater success in reducing excess weight^[17]. This observation is particularly relevant because most studies to date have relied on group-level analyses, which overlook patient-level heterogeneity and limit our understanding of the precise impact of microbiota on individual weight-loss trajectories. Similarly, microbiota-derived metabolites such as SCFAs have also been associated with surgical success. We previously identified a decrease in straight-chain SCFAs (such as acetate, valerate, and hexanoate) at 12 months post-surgery, which was associated with an improvement in glucose metabolism^[18].

Collectively, these observations suggest that postoperative weight-loss trajectories are shaped by complex interactions among clinical, metabolic, and microbial factors rather than by individual variables acting in isolation. Capturing such multidimensional relationships remains challenging when using conventional statistical approaches, particularly when dealing with high-dimensional microbiome data. In this context, machine-learning (ML) methods, especially those linked to explainability algorithms, as in the current work, offer a promising framework for integrating heterogeneous data sources and identifying patterns that may improve the prediction of individual responses to BS^[19].

Previous studies have applied ML approaches to predict weight-loss outcomes after BS using clinical, anthropometric, biochemical, psychological, or microbiome-derived variables. Similarly, presurgical gut microbiota profiles have been

associated with subsequent weight-loss response, although the available evidence remains inconsistent, with some studies reporting predictive microbial signatures and others finding limited predictive value for baseline microbiota alone^[20–22]. Beyond predictive performance, the clinical usefulness of ML models also depends on their interpretability. Explainability methods, such as SHapley Additive exPlanations (SHAP), can identify the variables contributing to model predictions at both the population and individual levels, potentially improving transparency and facilitating clinical interpretation^[23]. However, despite previous studies employing ML algorithms^[22,24], the use of explainable ML to jointly evaluate preoperative gut microbiota, circulating microbial metabolites, and detailed host characteristics in relation to both short- and mid-term weight-loss response after RYGB remains insufficiently explored.

Therefore, the objective of the present study was to integrate baseline family-level gut microbiota, circulating levels of SCFAs, and clinical and biochemical features to develop and internally validate models predicting weight-loss response at 12 and 24 months after RYGB. In addition, we used SHAP analyses to provide global and individual-level explanations of the ML predictions and to evaluate the potential complementary contribution of microbial, metabolic, and clinical information.

Methods

Study design and participants

This single-center study had a prospective cohort design. All participants underwent a standard preoperative assessment established by the hospital. Eligible participants were adults aged 18–65 years with severe obesity who were candidates for BS according to the 2022 ASMBS and IFSO consensus criteria^[25], with a documented history of severe obesity for ≥ 5 years and a minimum postoperative follow-up of 24 months. Exclusion criteria included major cardiovascular events within the previous 6 months, any acute inflammatory or acute infectious disease, previous BS, or a history of advanced renal or hepatic disease. To minimize baseline confounding of gut microbiota and SCFA profiles, participants reporting systemic antibiotic use within the previous 3 months and PPI or probiotic/prebiotic supplementation within 1 month prior to surgery were also excluded. All patients received standard perioperative antibiotic prophylaxis (2 g of cefazolin/20 minutes before the first incision) according to institutional surgical protocols; this exposure was uniform across all participants.

All study participants were scheduled to undergo RYGB surgery. Patients were followed up periodically in the outpatient setting, with clinical and anthropometric parameters updated at each visit, as previously described^[26]. Baseline data from all enrolled patients with available stool samples and complete clinical and anthropometric measurements were included ($n = 145$). During follow-up, two participants did not complete the evaluation; therefore, 143 patients were evaluated. Written informed consent was obtained from all participants after they received detailed information regarding the aims and procedures of the study. The study was conducted in accordance with the Declaration of Helsinki and approved by the Research Ethics Committee Hospital (information anonymized for review purposes). This study adhered to the STROBE guidelines for cohort studies. This manuscript complies with the TITAN Guidelines 2025^[27]. Artificial intelligence (AI) tools were used exclusively

for grammatical editing, language refinement, and assisting with code syntax debugging in R and Python. AI was not used to autonomously conduct data analysis, select statistical models, or interpret study results; all analytical decisions, code executions, and interpretations were performed solely by the authors.

Main outcomes

The primary outcome was the percentage of total weight loss (% TWL) at 12 and 24 months after RYGB. At each follow-up time point, participants who achieved %TWL $\geq 30\%$ were classified as high responders, whereas those achieving %TWL $< 30\%$ were classified as low responders. Response status was therefore defined separately at 12 and 24 months. The 30% cut-off was selected based on Spanish multicenter evidence indicating that an average %TWL of approximately 30% at 12 months after RYGB represents clinically significant weight loss^[28].

Anthropometric and other biochemical and clinical parameters

Anthropometric measurements, including height, weight, and waist circumference, were obtained using standardized procedures. Measurements were taken by trained professionals at the Service of Endocrinology and Nutrition of the HCUVA during the preoperative evaluation of patients with severe obesity.

Preoperative blood samples were collected after a 12-hour overnight fast, and the serum was separated by centrifugation. Serum levels of glucose, total cholesterol, high-density lipoprotein cholesterol (HDL-c), triglycerides, alanine aminotransferase (ALT), aspartate aminotransferase (AST), and gamma-glutamyl transferase were determined using standardized methods with a Cobas Analyzer c702 (Roche). Plasma HbA1c was measured using an HLC[®]-723G8 glycohemoglobin analyzer (Tosoh Bioscience), and serum insulin was quantified using a Cobas Analyzer e801 (Roche). Insulin resistance was estimated using the homeostatic model assessment (HOMA-IR) formula: insulin ($\mu\text{U/mL}$) $\times$ glucose (mmol/L)/22.5. Plasma protein and hormone concentrations were analyzed by ELISA.

Determination of plasma SCFA levels by LC-MS

The methodology used to evaluate SCFAs has been described previously^[18]. Briefly, eight SCFAs were analyzed: acetate, butyrate, heptanoate, hexanoate, isobutyrate, isovalerate, propionate, and valerate. An Agilent Technologies 6470 triple quadrupole mass spectrometer was operated with an electrospray ionization source in positive mode. For these experiments, 100 μL of plasma was treated with methanol and centrifuged at 14 000 rpm for 10 min at 4°C. The supernatant was recovered and incubated with water, O-benzylhydroxylamine hydrochloride, and N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride for 1 hour at room temperature. The samples were then extracted with dichloromethane, which was removed, and the residue was reconstituted with a mixture of methanol and water. Quantification was carried out using standards (Volatile Free Acid Mix #46975-U, Sigma-Aldrich, St. Louis, MO).

Gut microbiota analysis in fecal samples by 16S rRNA gene sequencing

Fecal samples were obtained at baseline (preoperative) and during follow-up (12 and 24 months post-RYGB). The gut microbiota

analysis was carried out at Microomics Systems S.L. (Barcelona, Spain). Baseline microbiota profiles were used for predictive modeling of postoperative weight loss outcomes. Feces were collected and preserved by healthcare staff using OMNIgene.GUT kits from DNA Genotek (Ottawa, ON, Canada) following the supplier's standard guidelines. Gut microbiota composition was assessed using 16S rRNA gene sequencing of the stool samples. Genomic DNA was extracted from 200 mg of fecal material using the QIAamp DNA Stool Mini Kit (Qiagen, Manchester, UK) following the manufacturer's instructions. The DNA concentration and integrity were evaluated using an Agilent 2200 TapeStation. For library preparation, DNA samples (5 ng/μL) were PCR-amplified using the 16S Metagenomics Kit (Thermo Fisher Scientific, Waltham, MA, USA), targeting bacterial 16S rRNA hypervariable regions V2–4–8 and V3–6, 7–9. The primer sequences were those included in the proprietary Ion 16S Metagenomics Kit. Amplicons were purified using AMPure XP beads (Beckman Coulter, Pasadena, CA, USA). Sequencing libraries were constructed using the Ion Plus Fragment Library Kit (Thermo Fisher Scientific) and Ion Xpress Barcode Adapters 1–16 (Thermo Fisher Scientific) to barcode purified amplicons. Libraries were purified using AMPure XP beads and quantified by fluorescence. Sequencing was performed on the Ion S5 platform.

Raw sequencing data were processed using QIIME2 2022.2. The median sequencing depth was 155 358.1 reads per sample, with an interquartile range of 109 632.0 to 182 101.2. Raw reads were demultiplexed by variable region using the Split_On_Primer.py script (GitHub: Y-Lammers/Split_on_Primer). Reads corresponding to each variable region were filtered, and the mean read lengths were 238.3 (V2), 186.7 (V3), 260 (V4), 241.7 (V6), 257.5 (V8), and 190.8 (V9). Sequences were processed using DADA2, with a chimera-removal procedure. Taxonomic assignment was performed using regional k-mer-based alignment to a reference, and then estimating relative abundances using a maximum-likelihood estimator against the SILVA database, version 138 (99% OTUs, full-length sequences)^[29,30]. Formal negative extraction or PCR controls and automated bioinformatic contaminant-filtering procedures (e.g., decontam) were not incorporated into the sequencing workflow; however, all samples were processed following identical standardized laboratory protocols.

The microbiome features were aggregated at the family level to reduce sparsity and improve robustness, as genus-level features derived from 16S rRNA sequencing are often highly sparse and subject to annotation variability, which may inflate dimensionality and reduce statistical power and model stability. Aggregation at the family level provides more consistently detected and reproducible features, thereby improving the signal-to-noise ratio and the reliability of downstream predictive modeling.

Statistical analysis

Data are presented as means $\pm$ standard deviations. Participants were classified separately at 12 and 24 months as high responders (%TWL $\geq 30\%$) or low responders (%TWL $< 30\%$), according to the predefined primary outcome, as described previously. Differences between groups were analyzed using Welch's *t*-test for clinical and biochemical parameters and the Wilcoxon rank-sum test for microbiota features. The false discovery rate (Benjamini-Hochberg) was calculated to control for multiple comparisons, and the effect size (Cohen's *d*) was reported. An analysis of covariance was also carried out to adjust for the effects of sex, age, and baseline

weight. The microbiota features were evaluated using the MicrobiomeAnalyst platform (<https://www.microbiomeanalyst.ca/>). Alpha diversity was evaluated using the Chao1 and Shannon indices, although only the former was reported, since there were no statistically significant differences between them. For exploratory and univariable microbiota analyses conducted using the MicrobiomeAnalyst platform, centered log-ratio (CLR) transformations were applied. For the predictive machine-learning framework, microbiome features consisted of family-level relative abundances that were centered and scaled (normalized) within each training fold without compositional transformations. This analytical strategy was selected for the predictive modeling workflow, but it should be interpreted with consideration of the inherently compositional nature of microbiome data.

For logistic regression (LR), three models were developed separately for the 12- and 24-month outcomes: (1) a microbiome-only model, (2) a clinical-metabolic model including clinical, anthropometric, biochemical, and circulating SCFA variables, and (3) a combined model including both predictor blocks. Model development and internal validation were performed in the full cohort of 143 participants using stratified repeated five-fold cross-validation with 20 repetitions. In each resampling iteration, 80% of participants were assigned to the analysis set and 20% to the assessment set. No independent hold-out test set was used. The same resampling splits were applied to all three models. Within each analysis fold, missing numeric values were imputed using the training-fold median, missing categorical values using the training-fold mode, novel categorical levels were handled before dummy-variable generation, zero-variance predictors were removed, and continuous predictors were standardized. These preprocessing parameters were estimated exclusively in the analysis fold and then applied to the corresponding assessment fold. LASSO-penalized LR was fitted using the *glmnet* library, and the penalty parameter was selected according to cross-validated area under curve (AUC). Model sparsity and predictor-selection stability were additionally evaluated across the 100 resampling iterations. For each model, we recorded the number of predictors assigned non-zero LASSO coefficients, the frequency with which each predictor was selected, its median non-zero coefficient, and the consistency of the coefficient direction.

Out-of-fold probabilities were generated for each assessment fold and averaged for each participant across the repeated resampling procedure. Candidate predictor sets were previously reduced using clinical plausibility, univariable associations, correlation screening, and variance-inflation-factor criteria. Because this candidate-variable preselection was performed before repeated cross-validation, the resulting internal performance estimates may retain some optimism. The complete workflow for data preprocessing, model fitting, internal validation, generation of out-of-fold predictions, and performance assessment is summarized in Supplemental Digital Content Figure S1, available at: <https://links.lww.com/JS9/H662>. To assess the incremental predictive value of microbiome information, the microbiome-only, clinical-metabolic, and combined models were evaluated using identical repeated cross-validation folds. Predictions were generated out-of-fold and averaged at the participant level across repetitions. AUCs were reported with 95% bootstrap confidence intervals, and the clinical-metabolic and combined models were compared using paired differences in AUC and DeLong tests. Incremental probabilistic accuracy was assessed using differences in Brier scores. Calibration was evaluated using calibration-in-the-large, calibration intercepts, calibration slopes, and calibration plots. Integrated discrimination improvement and

continuous net reclassification improvement were estimated as exploratory measures. Decision-curve analysis was used to compare the net benefit of the models across a range of threshold probabilities. All reported performance estimates, including AUC, Brier score, calibration measures, and decision-curve analyses, were calculated from participant-level out-of-fold predictions generated using stratified, repeated five-fold cross-validation with 20 repetitions. No separate hold-out test set was used.

Finally, logistic recalibration (intercept–slope adjustment) was applied to obtain recalibrated probabilities. Different ML analyses and SHAP value calculations were carried out using the SIBILA platform (the source code is available at <https://github.com/bio-hpc/sibila.git> accessed on 10 January 2026)^[31]. SHAP values were used exclusively to provide a descriptive interpretation of the fitted model by quantifying the contribution of each feature to individual predictions. SHAP values were not used for variable selection, statistical inference, or causal interpretation. Because SHAP rankings may be influenced by preprocessing, predictor correlation, model specification, and sample composition, they were interpreted as exploratory summaries of model behavior.

Results

Baseline clinical and metabolic characteristics of the study population

Table 1 presents the baseline anthropometric and clinical characteristics of the study participants who completed the 24-month

Table 1		
Baseline characteristics of the study participants.		
Variable	Total participants (n = 143)	95% CI (lower–upper)
Age (y)	46.5 ± 11.0	(44.7–48.3)
Sex (% women)	74.1%	
Weight (kg)	118.6 ± 21.6	(115.1–122.2)
BMI (kg/m ²)	43.4 ± 6.1	(42.4–44.4)
Waist (cm)	126.6 ± 14.1	(124.2–128.9)
Body fat (%)	44.7 ± 5.8	(43.8–45.7)
SBP (mmHg)	138.9 ± 19.1	(135.7–142.0)
DBP (mmHg)	84.7 ± 11.5	(82.8–86.6)
Cholesterol (mg/dL)	163.0 ± 31.6	(157.7–168.2)
Triglycerides (mg/dL)	198.0 ± 141.9	(174.5–221.6)
LDL-c (mg/dL)	86.7 ± 29.2	(81.7–91.7)
HDL-c (mg/dL)	42.1 ± 11.4	(40.2–44.0)
Glucose (mg/dL)	102.5 ± 34.0	(96.9–108.1)
Hb1Ac (%)	6.0 ± 1.2	(5.8–6.2)
ALT (U/L)	22.7 ± 13.1	(20.5–24.8)
AST (U/L)	20.8 ± 11.6	(18.8–22.8)
Creatinine (mg/dL)	0.7 ± 0.3	(0.7–0.7)
GFR _{CKD-EPI} (mL/min)	90.0 ± 14.1	(87.6–92.3)
Acetate (μmol/L)	42.5 ± 68.9	(30.6–54.4)
Butyrate (μmol/L)	1.1 ± 1.0	(0.9–1.2)
Heptanoate (μmol/L)	0.9 ± 0.6	(0.8–1.0)
Hexanoate (μmol/L)	2.3 ± 2.8	(1.9–2.8)
Isobutyrate (μmol/L)	0.4 ± 0.2	(0.3–0.4)
Isovalerate (μmol/L)	0.8 ± 0.8	(0.7–0.9)
Propionate (μmol/L)	4.8 ± 5.0	(4.0–5.7)
Valerate (μmol/L)	0.5 ± 0.3	(0.5–0.6)

Data are shown as mean ± SD and 95% confidence interval.

ALT, alanine transaminase; AST, aspartate aminotransferase; BMI, body mass index; DBP, diastolic blood pressure; GFR, glomerular filtration rate; HDL-c, high-density lipoprotein cholesterol; LDL-c, low-density lipoprotein cholesterol; SBP, systolic blood pressure.

follow-up period ($N = 143$). Most of the participants were women ($n = 106$, 75%). Consistent with the criteria for undergoing RYGB, the mean BMI was over 40, a situation mirrored by the high mean body fat percentage. Systolic blood pressure was slightly elevated, although diastolic blood pressure was within the normal range. Fasting glucose and HbA1c parameters were also within normal ranges. The patients showed a typical pattern of obesity-associated dyslipidemia, with marked and highly variable hypertriglyceridemia and low HDL levels, whereas LDL and total cholesterol levels were within moderate ranges. Furthermore, there were no evident signs of significant renal impairment at baseline, and mean renal function was preserved (considering creatinine and glomerular filtration rate). Plasma SCFA concentrations were also variable, potentially reflecting interindividual differences in gut microbiota composition (Table 1). Additional baseline parameters related to hepatic function, systemic inflammation, and hormonal profiles are shown in Supplemental Digital Content Table S1, available at: <https://links.lww.com/JS9/H662>. Baseline anthropometric, clinical, biochemical, and circulating SCFA characteristics according to 12-month response status are reported in Supplemental Digital Content Table S2, available at: <https://links.lww.com/JS9/H662>.

Changes in gut microbiota, clinical, and metabolomic features according to postoperative weight-loss response and follow-up

To investigate whether baseline gut microbiota composition, circulating SCFAs, and clinical and metabolomic features differed according to postoperative weight-loss response, participants were stratified into high-responders (%TWL $\geq 30\%$) and low-responders (%TWL $< 30\%$) at 12 and 24 months after RYGB.

Regarding gut microbiota diversity, individuals with greater weight loss at 12 months (Fig. 1A) and 24 months (Fig. 1B) were characterized by higher diversity than those who did not reach 30% TWL from baseline ($P = 0.009$ and $P = 0.021$, respectively). These differences remained statistically significant even after adjusting for sex and clinical antecedents at 12 months post-surgery, but not at 24 months ($P_{adj} = 0.032$ and $P_{adj} = 0.060$, respectively). Conversely, no changes in microbiota diversity were observed at 12 months after surgery, either in the total population (Fig. 1C) or when patients were categorized as high responders or low responders according to the 30% TWL (Fig. 1D).

Figure 2 illustrates the differences in the gut microbiota composition at 12 and 24 months after the intervention. At 12 months post-surgery, several bacterial families showed significant differences in relative abundance between high- and low-responders, suggesting the presence of modest yet potentially relevant microbial signatures associated with early weight loss response (Fig. 2A). In contrast, stratification according to 24-month weight loss outcomes revealed more pronounced differences. At the later time point, a larger number of bacterial families reached statistical significance, with several families displaying stronger effect sizes (Fig. 2B). Notably, bacterial families such as *Beijerinckiaceae* and *Lactobacillaceae* were significantly associated with high-response status (Fig. 2, Supplemental Digital Content Figures S2 and S3, available at: <https://links.lww.com/JS9/H662>). Overall, these findings suggest that the baseline gut microbiota composition is more strongly related to sustained, long-term weight loss responses than to short-term outcomes, indicating that specific baseline microbial families were associated with response status at 12 and 24 months.

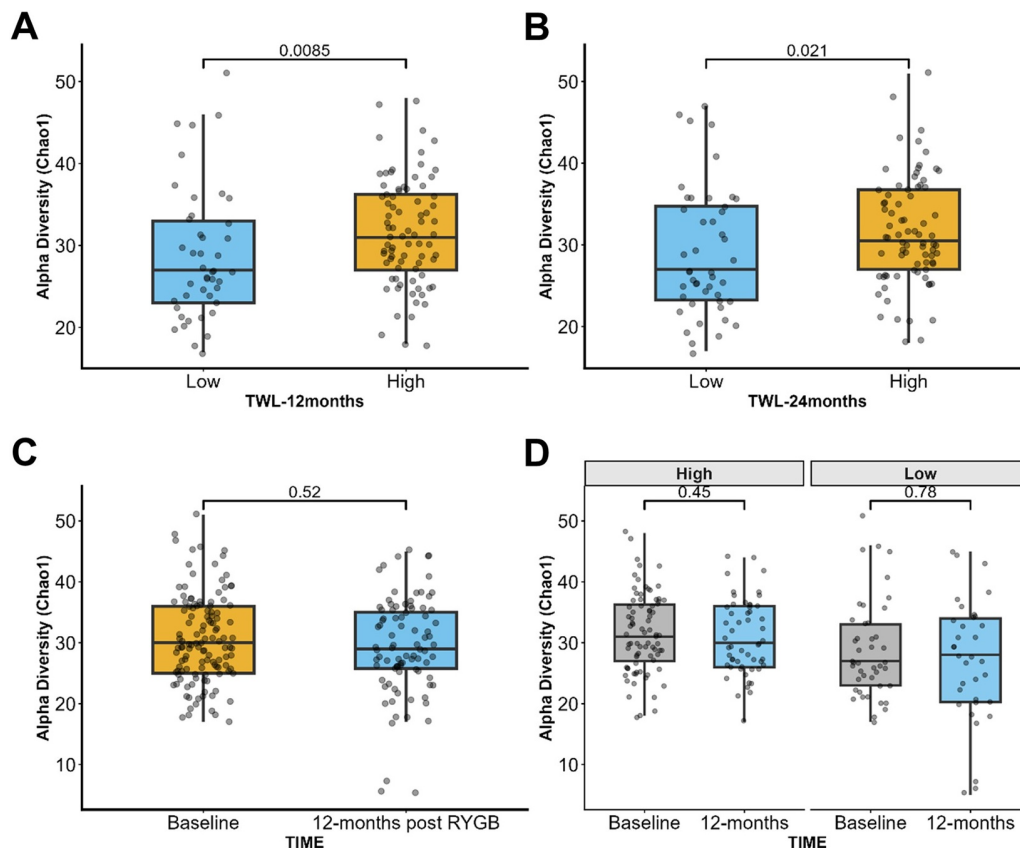


Figure 1. Alpha diversity (Chao1 index) of the gut microbiota according to weight loss response and follow-up time after RYGB. Alpha diversity was compared between participants with high-response status (%TWL $\geq 30\%$) versus low-response status (%TWL $< 30\%$) at 12 months (A) and 24 months (B). Longitudinal changes in alpha diversity were evaluated at 12 months post-RYGB in the entire study population (C) and in the participants stratified by weight loss response at 12 months (D). Box plots represent the median and interquartile range (IQR), and whiskers represent the $1.5 \times$ IQR; dots indicate individual observations. P-values were calculated using the Wilcoxon test. TWL, total weight loss; RYGB, Roux-en-Y gastric bypass.

Several differences were observed in the anthropometric, clinical, and metabolic parameters. As shown in Supplemental Digital Content Table S2, available at: <https://links.lww.com/JS9/H662>, high responders at 12 months were younger and had higher body weight and body fat mass, whereas fat-free mass was comparable between the groups; no association between sex and weight-loss response was observed. Notably, most of these differences were not statistically significant at 24 months (Supplemental Digital Content Table S3, available at: <https://links.lww.com/JS9/H662>).

Key predictors of weight loss at 12 months and 24 months post-surgery

Predictive models were developed to identify factors associated with weight loss at 12 months (Fig. 3) and 24 months (Fig. 4) after surgery. At 12 months, *Peptostreptococcaceae* and *Clostridiaceae* were the gut microbiota families most strongly associated with high-response status at 12 months, with odds ratios (ORs) of 4.9 and 2.7, respectively. In contrast, *Erysipelotrichaceae* and *Oscillospiraceae* were associated with a lower probability of significant weight loss (Fig. 3A). Among the clinical and metabolic parameters, C-reactive protein (CRP) and HbA1c showed the highest ORs for achieving %TWL $\geq 30\%$, while heptanoate and isobutyrate were related to lower weight loss after 12 months

(Fig. 3B). The combined model integrating the microbiome and clinical parameters also identified *Peptostreptococcaceae* as the feature with the largest model-associated estimate (Fig. 3C). However, the overall performance of this combined model was lower than that of the clinical-metabolic model (Fig. 3D).

Using paired out-of-fold predictions from repeated five-fold cross-validation, the clinical-metabolic and combined models achieved AUCs of 0.720 (95% CI 0.632–0.803) and 0.717 (95% CI 0.634–0.797), respectively, at 12 months. The incremental difference was negligible and not statistically significant (Δ AUC = -0.003 ; 95% bootstrap CI -0.070 to 0.067 ; paired DeLong $P = 0.917$). At 24 months, the corresponding AUCs were 0.676 (95% CI 0.585–0.763) and 0.731 (95% CI 0.637–0.816), respectively. Although the combined model showed numerically higher discrimination, the incremental difference was not statistically significant (Δ AUC = 0.055 ; 95% bootstrap CI -0.015 to 0.128 ; paired DeLong $P = 0.122$). Complete estimates of discrimination, probabilistic accuracy, calibration, reclassification, and net benefit are provided in Supplemental Digital Content Table S4, available at: <https://links.lww.com/JS9/H662> and Supplemental Digital Content Figure S4, available at: <https://links.lww.com/JS9/H662>. LASSO retained most of the previously reduced candidate predictor sets. Across repeated resampling, the median number of selected predictors ranged from 7 to 25, depending on the follow-up time and

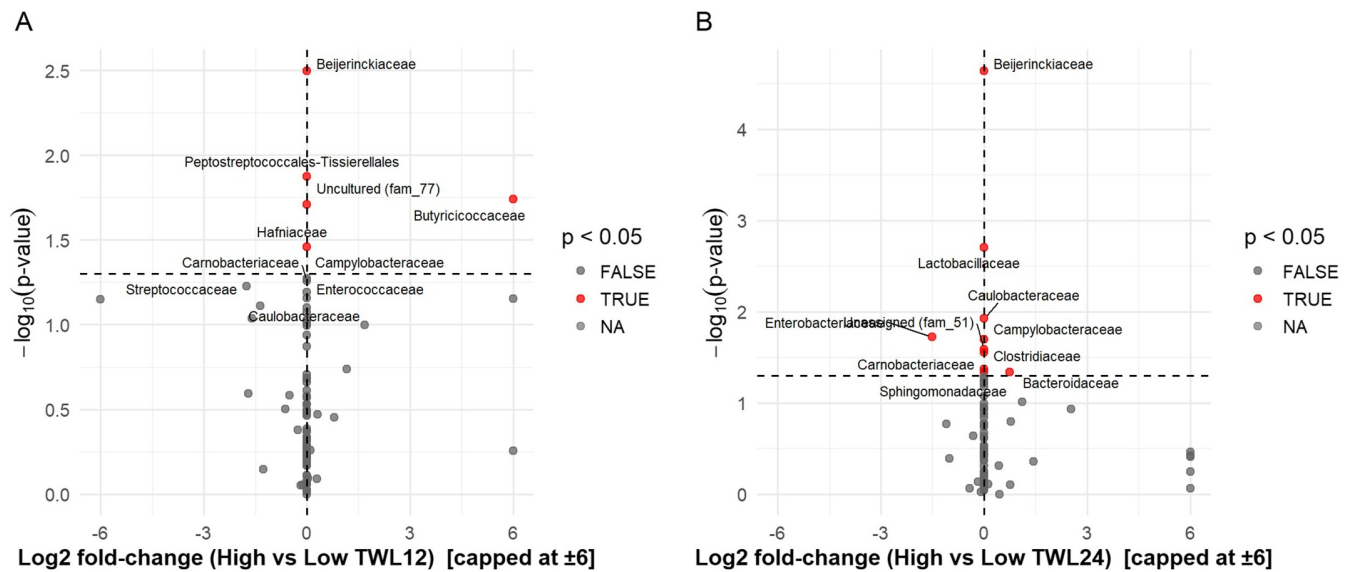


Figure 2. Baseline gut microbiota differences associated with weight loss response at 12 and 24 months after RYGB. Volcano plots depict differences in baseline gut microbiota family-level relative abundances between participants classified as high- versus low-responders at 12 months (A) and 24 months (B). The x-axis indicates the $\log_2$ fold-change in relative abundance (limited to ± 6 for visualization purposes). The y-axis represents statistical significance ($-\log_{10}(P\text{-value})$). Each point corresponds to a family. Red points indicate significant differences ($P < 0.05$), while gray points represent non-significant associations. Bacterial families with the lowest P -values were labeled.

predictor block, and predictor-selection frequencies were generally $\geq 95\%$. Nevertheless, some predictors showed less consistent coefficient directions across resamples. Complete model-sparsity and predictor-stability results are provided in Supplemental Digital Content Table S5, available at: <https://links.lww.com/JS9/H662>.

At 24 months, the prediction was more difficult and less precise, with only a limited number of gut microbiota families showing significant associations with weight loss outcomes. Nevertheless, consistent with the 12-month analysis, *Oscillospiraceae* remained predictive of a lower likelihood of achieving $\%TWL \geq 30\%$ (Fig. 4A). Among the clinical parameters, hemoglobin and folate exhibited the highest ORs; however, the association with hemoglobin did not reach statistical significance (Fig. 4B). The combined model integrating the microbiome and clinical parameters identified haptoglobin as the factor most strongly associated with high-response status at 24 months (Fig. 4C). Nonetheless, as observed at 12 months, the clinical-metabolic model demonstrated superior predictive performance compared to the combined model (Fig. 4D).

Across models, calibration slopes derived from out-of-fold predictions ranged from 0.308 to 0.463 (Supplemental Digital Content Table S4, available at: <https://links.lww.com/JS9/H662>), reflecting moderate overconfidence in predicted probabilities (slopes < 1). Brier scores ranged from 0.209 to 0.230 across time points and model types. For instance, at 12 months, the combined model achieved a calibration slope of 0.334, a calibration intercept of 0.316, and a Brier score of 0.225 (Supplemental Digital Content Table S4, available at: <https://links.lww.com/JS9/H662> and Supplemental Digital Content Figure S4, available at: <https://links.lww.com/JS9/H662>).

Because the underlying relationships between metabolic markers, microbiota features, and weight loss outcomes may be non-linear and interactive, conventional LR was complemented by an

ML approach. To enhance interpretability and clinical applicability, SHAP was used to provide both global and individual explanations of the model predictions. Among the evaluated algorithms, XGBoost yielded the best predictive performance (Supplemental Digital Content Tables S6 and S7, available at: <https://links.lww.com/JS9/H662>) and was therefore chosen for SHAP-based explainability analyses. These determinations were performed at the 12-month time point, as the model sensitivity and specificity were significantly reduced at 24 months, limiting their clinical relevance. The SHAP Beeswarm plot (Fig. 5A) summarizes the contribution of the most influential characteristics across all participants. Biochemical variables associated with inflammation, such as cholinesterase and CRP; SCFAs, including heptanoate; gut microbiota families, such as *Coriobacteriaceae*; and clinical and anthropometric parameters (ALT, uric acid, and leg perimeter) provided the greatest explanatory power for the model. Waterfall plots (Fig. 5B and 5C) illustrate the generation of individual predictions. In the participants classified as high-responders (probability 92.59%; Figure 5B), positive contributions from metabolites (isobutyrate, heptanoate, and butyrate) and electrolytes (sodium and chloride) shifted the prediction toward high-responders, while other variables (e.g., triglycerides and SHBG) contributed negatively. Conversely, in participants classified as low responders (predicted probability for high weight loss: 1.16%; Figure 5C), strong negative contributions from chloride and sodium, together with other variables, markedly reduced the predicted probability of achieving high weight loss. SHAP analysis highlighted the features that contributed most strongly to the predictions of the fitted model. These rankings describe model behavior and should not be interpreted as evidence of causality, biological mechanisms, or independent predictor effects.

Overall, these results suggest that the integrated profile of metabolic, clinical, and microbiota markers allows for better

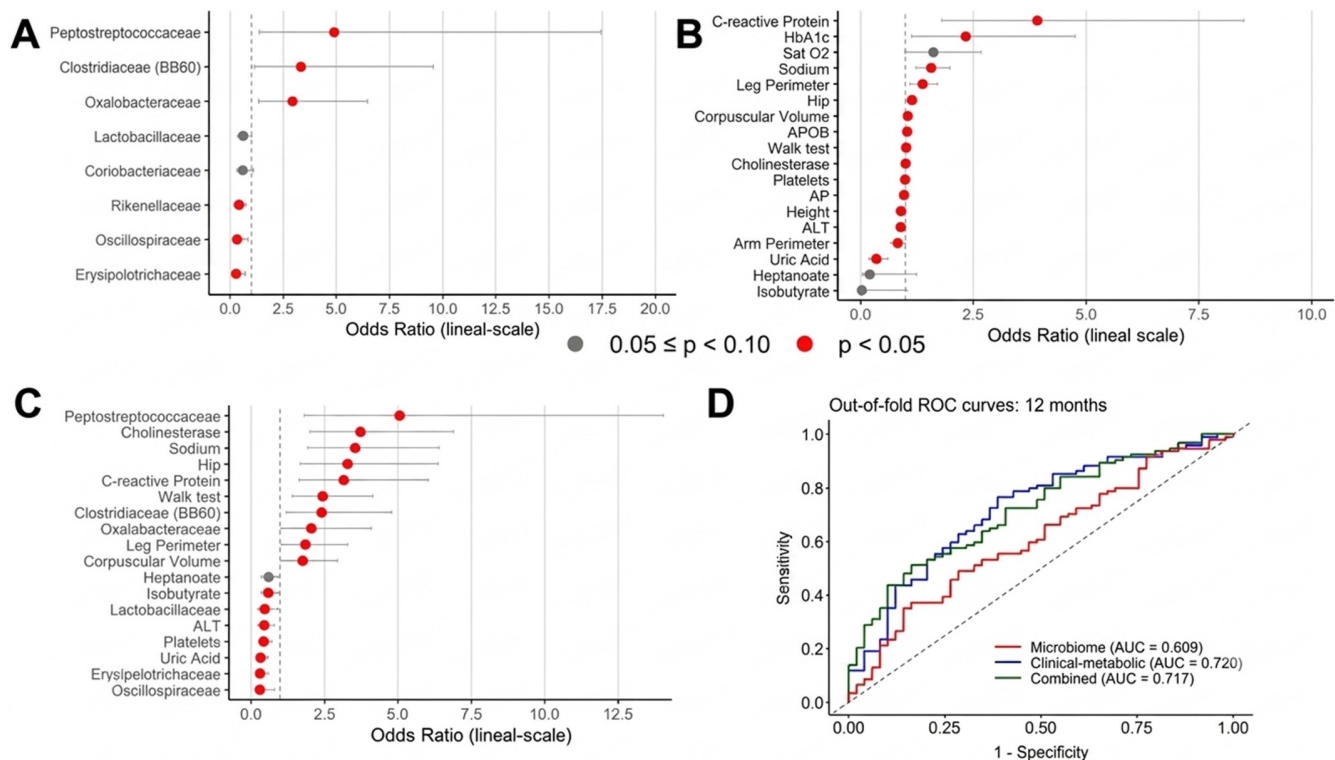


Figure 3. Baseline microbiome, metabolomic, and clinical features associated with high-response status at 12 months after RYGB. Panel A shows ORs and 95% confidence intervals for baseline gut microbiota families associated with high-response status. Panel B shows ORs and 95% confidence intervals for baseline clinical, anthropometric, biochemical, and circulating short-chain fatty acid variables associated with high-response status. Panel C displays the corresponding estimates for the combined microbiome, metabolomic, and clinical predictors. The vertical dashed line indicates the null value (OR = 1). Red points indicate $P < 0.05$, whereas gray points indicate $0.05 \leq P < 0.10$. These estimates describe the magnitude and direction of the observed associations and should not be interpreted as causal or mechanistic effects. Panel D shows participant-level out-of-fold receiver operating characteristic curves for the microbiome-only model (red), clinical-metabolic model (blue), and combined model (green), obtained using stratified, repeated five-fold cross-validation. ALT, alanine aminotransferase; AP, alkaline phosphatase; APOB, apolipoprotein B; AUC, area under the receiver operating characteristic curve; CI, confidence interval; CRP, C-reactive protein; HbA1c, glycated hemoglobin; OR, odds ratio; ROC, receiver operating characteristic; Sat O₂, oxygen saturation.

interpretation of the algorithm's decisions. The final model was implemented in a web-based proof-of-concept interface to facilitate the visualization of predictions and SHAP explanations. The app is for research and demonstration purposes only and should not be used for clinical decisions, and its availability does not represent external validation (available at https://huggingface.co/spaces/juanjo-ucam/12months_weightlossprediction, accessed on 28 December 2025).

Discussion

This study evaluated whether baseline family-level gut microbiota features, circulating SCFAs, and clinical-metabolic variables contributed to the prediction of high-response status after RYGB in patients with severe obesity at 12 and 24 months post-surgery, using the %TWL (greater than or less than 30%) as the primary outcome. Baseline diversity and specific family-level signatures, such as the abundance of *Peptostreptococcaceae* and *Prevotellaceae*, were associated with a high-response status after RYGB. The combination of clinical and metabolomic (SCFAs) variables enabled the development of internally validated exploratory prediction models, with the strongest discrimination observed at 12 months.

A key issue that deserves consideration is the differences in baseline microbiota diversity among the study groups. At both 12 and 24 months, patients with greater weight loss were characterized by greater microbiota diversity. While many previous studies have focused on microbiota changes after surgery, only a limited number have evaluated the contribution of the gut microbiome to the success rate of RYGB^[32]. A previous cross-sectional study observed no differences between those with a successful intervention and those with primary failure; however, this study was conducted in a very small group of patients ($n = 24$). Therefore, this lack of difference may be due to reduced statistical power^[21]. Another study conducted on a very small sample group ($n = 20$) observed a correlation between baseline diversity and weight loss^[33], which is in line with our observations. A common observation is that people with normal weight have higher diversity than people with obesity^[34] and that people with obesity have worse metabolic profiles^[35], which is consistent with the hypothesis that this greater baseline diversity may indicate a better response after the intervention.

Several factors may explain why higher diversity is associated with greater weight loss after RYGB. In diverse microbiota, many different species can perform similar functions (e.g., fiber fermentation and vitamin synthesis), and if one microbiome feature is lost (e.g., due to antibiotics or diet change), others can compensate

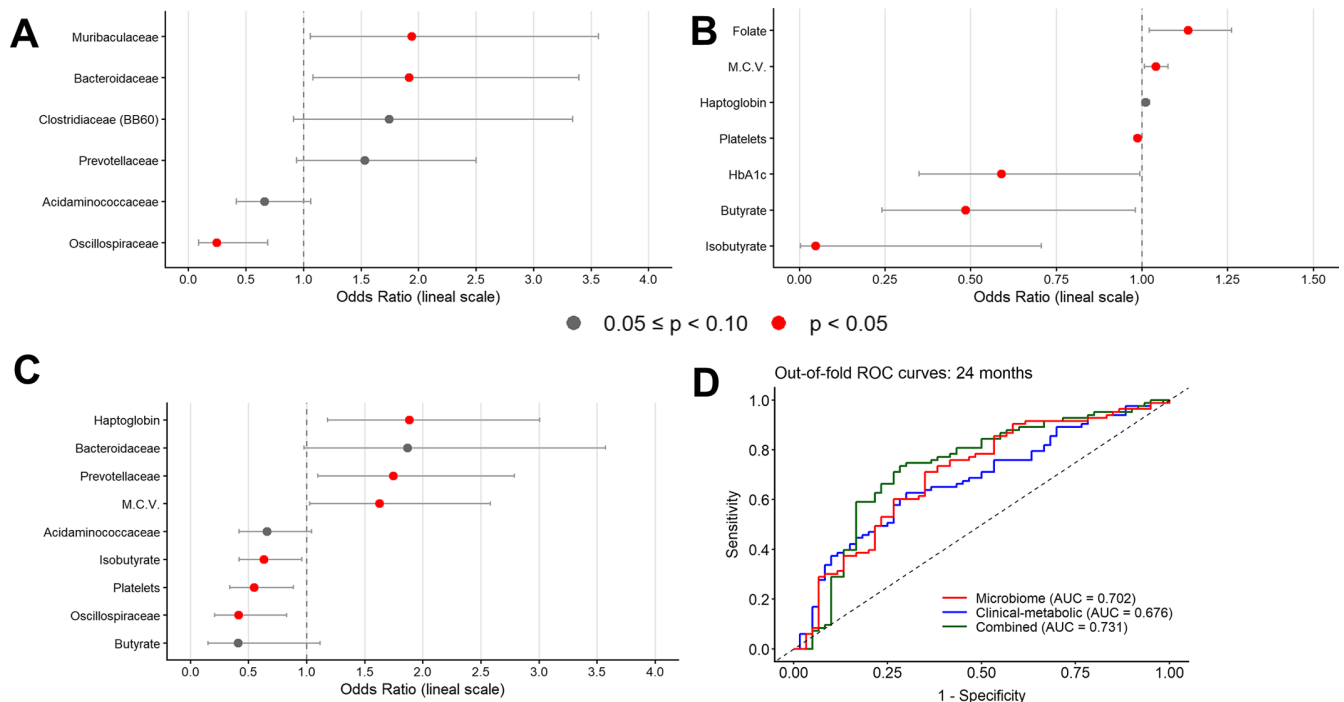


Figure 4. Baseline microbiome and clinical-metabolic features associated with high-response status at 24 months after RYGB. Panel A shows ORs and 95% confidence intervals for baseline gut microbiota families associated with high-response status. Panel B shows ORs and 95% confidence intervals for baseline clinical, anthropometric, biochemical, and circulating short-chain fatty acid variables associated with high-response status. Panel C displays the corresponding estimates for the combined microbiome, metabolomic, and clinical predictors. The vertical dashed line indicates the null value (OR = 1). Red points indicate $P < 0.05$, whereas gray points indicate $0.05 \leq P < 0.10$. These estimates describe the magnitude and direction of the observed associations and should not be interpreted as causal or mechanistic effects. Panel D shows participant-level out-of-fold receiver operating characteristic curves for the microbiome-only model (red), clinical-metabolic model (blue), and combined model (green), obtained using stratified repeated five-fold cross-validation. AUC, area under the receiver operating characteristic curve; CI, confidence interval; HbA1c, glycated hemoglobin; MCV, mean corpuscular volume; OR, odds ratio; ROC, receiver operating characteristic.

for it and maintain key metabolic outputs^[36]. Microbiota diversity supports host metabolic flexibility: a better capacity to adapt to changes in diet or energy balance, smoother glycemic responses, and more efficient energy harvesting without chronic overproduction of pro-inflammatory metabolites^[37].

Previous follow-up studies and systematic reviews have reported an increase in gut microbiome diversity after surgery^[11,38]. Conversely, we did not observe any longitudinal changes in the diversity of these patients. Several factors may underlie this issue, such as a shorter follow-up (1 or 3 months post-surgery) or smaller sample sizes^[11,39]. Interestingly, although we did not observe statistically significant changes, we detected an increase in diversity among participants with lower baseline diversity. This pattern suggests heterogeneity in responses: individuals starting with high baseline diversity showed little or no change over time, whereas those with lower baseline diversity tended to improve. These findings indicate that group averages may mask individual trajectories and support the need for a more detailed analysis.

Given that the existing literature has reviewed global microbiota changes after BS, our study evaluated whether baseline family-level microbiota features contributed to predicting post-operative response status. In the context of a moderate sample size and the limited taxonomic resolution of 16S rRNA sequencing, genus- and species-level features are often sparse and prone to annotation variability, which inflates dimensionality and penalizes power after multiple-testing corrections. Therefore, we aggregated reads at the family level to obtain more robust and

reproducible signals for predictive modeling and preoperative stratification. This approach allowed the exploratory evaluation of family-level microbiome features within an individualized prediction framework among patients undergoing RYGB.

In this context, specific microbiota families, such as *Beijerinckiaceae* and *Lactobacillaceae*, showed greater differences depending on weight-loss success. A greater abundance of *Beijerinckiaceae* was observed in those with greater weight loss at 12 and 24 months. To our knowledge, no previous reports have found differences in the abundance of this family in the context of BS, although a recent study found that a lower proportion of *Beijerinckiaceae* in people with obesity was associated with a higher production of inflammatory cytokines^[40]. Much information is available regarding the relationship between obesity and *Lactobacillaceae*. A large cohort study found that obesity is characterized by an increased abundance of Bacilli, mainly *Streptococcaceae* and *Lactobacillaceae*^[41]. Moreover, a previous study described obesity-associated gut microbiota enriched in *Lactobacillus*, although the authors also noted that some species were linked to more favorable metabolic profiles^[42].

Focusing on prediction ability, specific bacterial taxa have been identified as candidate predictors or correlates of postoperative weight-loss response. Higher pre- or postsurgical abundances of *Akkermansia*, *Bacteroides*, and certain *Proteobacteria*, such as *Escherichia coli*, are associated with better weight-loss outcomes^[43–45]. In fact, in this study, the higher abundance of

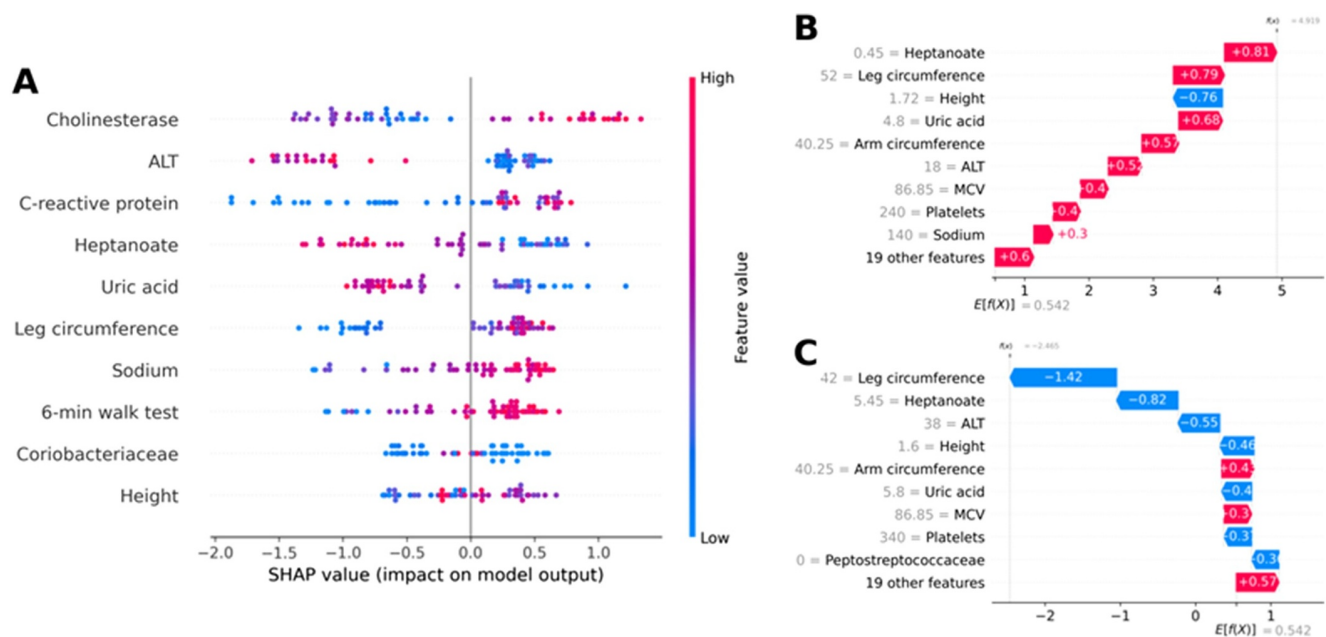


Figure 5. SHAP-based interpretation of the model predicting total weight loss at 12 months post-surgery. (A) Beeswarm plot of SHAP values derived from the XGBoost model for the dependent variable, the percentage of total weight loss (%TWL) at 12 months (high responder: %TWL $\geq 30\%$; low responder: %TWL $< 30\%$). The x-axis represents SHAP values, indicating the contribution of each variable to the model prediction, while the color scale reflects the value of the feature (blue = low value; pink = high value). The variables are ordered by overall importance. (B) Example of a SHAP waterfall plot for a participant classified as a high responder (probability of a high response = 92.59%). Bars represent the contribution of variables with the greatest impact on the model's prediction value: positive values (red) increase the probability of a high response, and negative values (blue) decrease it. (C) SHAP waterfall plot for a participant classified as a low-responder (probability of a high response = 16%). This interpretation is analogous to (B). ALT, alanine aminotransferase; $E[f(X)]$, expected model output; $f(x)$, model output for the individual prediction; MCV, mean corpuscular volume; SHAP, SHapley Additive exPlanations.

Proteobacteria was associated with greater weight loss, especially in the long term. Conversely, higher presurgical levels of *Bacteroidaceae*, *Bacteroides uniformis*, and Firmicutes were linked to poorer weight loss^[46].

The present findings should be interpreted within an emerging but heterogeneous literature on microbiome- and ML-based prediction after BS. Ben Izhak *et al* reported that presurgical gut microbiome information could predict subsequent weight-loss outcomes^[47], whereas Gutiérrez-Repiso *et al* identified distinct microbial signatures in patients experiencing successful weight loss, primary failure, or weight regain after BS^[21]. In contrast, Salazar *et al* found that baseline gut microbiota was not predictive of excess weight loss^[20]. These conflicting findings suggest that the predictive value of the preoperative microbiome remains uncertain and may depend on the type of surgical procedure, the definition of treatment success, the duration of follow-up, the sequencing and bioinformatic methods used, sample size, and whether microbiome information is evaluated alone or in combination with host and metabolic variables.

ML models based on clinical, anthropometric, biochemical, and psychological characteristics have also been developed to predict weight-loss success after RYGB or sleeve gastrectomy^[20,45]. Furthermore, previous studies combining microbiome data and ML have identified reproducible microbial signatures associated with RYGB, while longitudinal metagenomic and metabolomic studies have characterized substantial postoperative changes in microbial composition and metabolites^[48]. Therefore, the application of microbiome analysis or ML to bariatric-surgery outcomes is not, by itself, novel.

The contribution of the present study lies instead in the integration of exclusively preoperative family-level microbiota, circulating SCFAs, and detailed clinical and biochemical variables within the same prospective RYGB cohort, with outcomes assessed at both 12 and 24 months. An additional distinctive aspect is the use of explainable ML through SHAP analyses, which allows us to examine not only overall feature importance but also the contribution of individual variables to patient-specific predictions. This individual-level interpretability may improve the transparency of the models and generate clinically meaningful hypotheses, although it does not establish causal relationships or guarantee clinical utility.

Importantly, the clinical-metabolic models showed similar or superior predictive performance compared with the combined models in some analyses. Therefore, our findings do not demonstrate that microbiome measurement currently improves prediction beyond routinely available clinical information. Rather, these variables may provide complementary biological and interpretive information, whose incremental predictive value should be evaluated in independent multicenter cohorts.

From a translational perspective, the present findings highlight a critical distinction among predictive power, practical feasibility, and clinical utility. While baseline gut microbiota features offer physiological insights, the clinical-metabolic model achieved predictive performance comparable to the combined model. Clinical and biochemical variables are routinely collected as part of standard preoperative bariatric evaluations and incur minimal additional costs, have rapid turnaround times, and require no specialized infrastructure. In contrast,

targeted microbiome profiling entails high-throughput sequencing and specialized bioinformatic pipelines, has turnaround times of several weeks, and is hampered by a current lack of standardized clinical reference ranges. Therefore, while standard clinical–metabolic metrics represent the most pragmatic and scalable approach for routine preoperative risk stratification, baseline microbiome features remain key for patient-specific profiling. Rather than being redundant, microbiome profiling provides unique pathophysiological signatures that may help identify distinct patient sub-phenotypes, ultimately paving the way for individualized post-surgical monitoring and targeted metabolic interventions.

A key aspect of prediction studies, such as the present study, is the sensitivity and specificity of the models. Both the LR and ML models achieved useful performance at 12 months, although they were lower at 24 months. Furthermore, SHAP values provided individual-level descriptions of model behavior and improved interpretability. However, these explanations do not establish clinical usefulness or readiness for implementation. It is important to note that ML is complementary to LR and is not a substitute. In this regard, some studies have reported that baseline microbiota composition alone is insufficient for reliable prediction, emphasizing the need for integrated models including genetic, behavioral, and metabolic factors^[20,46], as implemented in the present study. However, adding microbiome features did not consistently improve predictive performance. Clinical and metabolomic variables may provide more direct and stable information on postoperative weight-loss response, whereas baseline microbiome profiles are affected by substantial interindividual variability. Residual variability related to diet, medication use, bowel habits, and other host factors may have further attenuated their incremental predictive value. Therefore, the performance of the clinical–metabolic model should not be interpreted as evidence that the microbiome is biologically unimportant, but rather that the baseline microbiome features in this cohort did not consistently improve individual-level prediction. In this regard, several authors have observed that microbial metabolites, such as the production of SCFAs, secondary bile acids, and specific amino acid fermentation products, are better predictors of post-surgical weight loss and metabolic health than specific microbial features^[49,50].

In the LR model, *Peptostreptococcaceae* abundance, cholinesterase, sodium, hip circumference, and CRP were identified as model-associated features related to high-response status. In line with this observation, in adults with severe obesity undergoing BS, a longitudinal study showed that *Peptostreptococcaceae* abundance decreased within 3 months after surgery, in parallel with metabolic improvements, suggesting that its higher preoperative levels are part of the severe obesity dysbiosis pattern^[51]. A review on obesity-related dysbiosis highlighted *Peptostreptococcaceae* among the Firmicutes families linked to proinflammatory states and higher BMI^[52]. The higher abundance of these bacteria may be associated with higher CRP and cholinesterase levels. In fact, in abdominal obesity and metabolic syndrome, higher serum butyrylcholinesterase activity has been correlated with higher BMI, insulin resistance, triglyceride levels, and other cardiometabolic risk markers, consistent with a state of low-grade systemic inflammation^[53]. ML and SHAP analyses also identified *Coriobacteriaceae* as a feature contributing to model predictions of postsurgical weight-loss response. The relationship between obesity and the abundance of this family in the gut microbiota has been previously described; however, its

clinical relationship is heterogeneous. While some studies have associated its abundance with a better metabolic profile^[54], others have associated it with metabolic impairment^[17]. Probably, as Kim *et al* suggest, this relative abundance, both in metabolically “better” and “worse” phenotypes, may depend on the dominant genera and host context^[55], which reinforces the need to interpret the role of the microbiota globally, along with a proficient clinical and metabolic characterization, as is the case in the present study. Moreover, because the microbiota and SCFA profiles used in the ML models were obtained preoperatively, these associations are less likely to reflect reverse causation from postoperative weight loss or dietary changes.

In the long term (24 months), the *Prevotellaceae* family emerged as the most influential feature at 24 months. Multiple studies have associated *Prevotella*-dominated communities with obesity, including severe obesity, although their role is complex and strongly diet-dependent. In a previous study conducted in a similar cohort, *Prevotellaceae* were part of the baseline signature of “responders” (those with better weight loss) compared with non-responders, suggesting that *Prevotella* may predict a more favorable response to surgery or diet, as observed in our study^[17]. However, the low accuracy of these models hampers data interpretation.

In a clinical context, our findings support further evaluation of preoperative microbiome profiling as an exploratory source of complementary information for weight-loss prediction after RYGB to move toward more personalized care. Interestingly, several of the strongest predictors (e.g., CRP and cholinesterase, together with cardiometabolic markers such as triglycerides and blood pressure) point to a more proinflammatory, metabolically dysregulated profile among those who achieved greater weight loss. This apparent paradox may reflect that patients with higher inflammatory levels and worse metabolic profiles have more room for improvement and experience a larger absolute metabolic shift after surgery, resulting in greater weight loss. If validated in external cohorts, these signatures could be integrated into preoperative counseling programs to maximize long-term outcomes. Moreover, SHAP values were used to improve the transparency and interpretability of the predictive models by identifying the features that most strongly influenced the model outputs at the individual and population levels. These explanations are model-specific and should be interpreted as hypothesis-generating rather than evidence of causal biological mechanisms.

Several limitations should be considered when interpreting these findings. First, although repeated cross-validation was used to obtain internally validated performance estimates, the models were developed and evaluated within a single-center cohort and were not externally validated; therefore, their generalizability to other populations, healthcare settings, sequencing platforms, and surgical practices remains uncertain. Second, microbiota profiling relied on 16S rRNA gene sequencing, which provides limited taxonomic resolution compared with shotgun metagenomics. In addition, while CLR transformations were used for exploratory analyses, predictive models were developed using relative abundances without compositional data transformations, which may introduce compositionality artifacts. While this approach is commonly used in predictive modeling studies, microbiome sequencing data are inherently compositional, and alternative analytical frameworks may provide complementary information and should

be explored in future studies. Furthermore, negative extraction controls and automated contaminant-filtering procedures were not incorporated into the sequencing workflow. Although all samples were processed under standardized single-batch protocols, the potential presence of low-level background contaminants cannot be fully ruled out. Third, although major microbiome-modifying exposures were minimized through exclusion criteria (e.g., antibiotics, PPIs, and probiotics), residual confounding related to diet, medication use, bowel habits, and other host factors cannot be fully excluded. Fourth, predictive performance was lower at 24 months, likely reflecting increased interindividual heterogeneity over time and the lower proportion of patients achieving marked weight loss at this time point. Fifth, as in any predictive-modeling study, overfitting and optimistic performance estimation are potential concerns. To mitigate these risks, we used repeated cross-validation with out-of-fold performance estimation and implemented preprocessing exclusively within the training data during each resampling iteration. Nevertheless, candidate predictor reduction was performed before repeated cross-validation, which may have introduced a degree of optimism into the reported performance estimates. Finally, although the clinical-metabolic models performed similarly to or slightly better than the combined models in some analyses, this should not be interpreted as evidence that the microbiome is biologically unimportant; rather, baseline microbiota features did not consistently improve individual-level prediction in this cohort. Consequently, the models should be regarded as exploratory and internally validated, and their clinical utility requires confirmation through external validation in independent multicenter cohorts.

Importantly, the associations identified in this study should not be interpreted as causal. Although baseline microbiota and SCFA profiles were used to predict postoperative outcomes, our modeling approach was observational and does not allow causal inference. The identified microbial and metabolomic features should, therefore, be considered as predictive biomarkers rather than mechanistic drivers of weight-loss response. Future longitudinal and interventional studies should focus on multicenter validation, higher-resolution multi-omics approaches, and mechanistic research to strengthen translation and improve the clinical applicability of these models.

In conclusion, our findings indicate that baseline gut microbiota profiles summarized at the family level by 16S rRNA sequencing, together with host metabolic and inflammatory features, were associated with postoperative response and provided complementary information to internally validated prediction models after RYGB, particularly at 12 months. Beyond descriptive postoperative changes, this study highlights the importance of complementing microbiome analyses (including bacterial metabolites) and establishing clinical markers to support preoperative stratification, personalized counseling, and follow-up. Although predictive performance was attenuated at 24 months, taken together, conventional LR and explainable ML indicate that postoperative weight-loss outcomes are not entirely unpredictable; rather, they are partly associated with a measurable preoperative host-microbiome profile that may support further research on individualized risk estimation. If confirmed in external cohorts and with higher-resolution multi-omic profiling, these signatures could be evaluated for potential use in stratification in externally validated studies, toward a more precise and clinically reliable framework.

Ethical approval

The study protocol was approved by the Ethics and Clinical Research Committee of Virgen de la Arrixaca University Hospital on 31 March 2020 (Reference number: 2020-2-4-HCUVA).

Patient or volunteer consent

Informed consent was obtained from all participants prior to their inclusion in the study.

Sources of funding

This research was funded by the Institute of Health “Carlos III” (ISCIII) through project PI23/00171 and co-funded by the European Union (grant numbers PI20/00505 and PI23/00171). I.R.-M. is supported by a Rio Hortega contract from the ISCIII (CM24/00019, ISCIII, Spain; co-funded by the European Commission). M.A.M-S is supported by a PFIS predoctoral fellowship from the ISCIII (FI21/00003, ISCIII, Spain; co-funded by the Fondo Europeo de Desarrollo Regional-FEDER).

Author contributions

Conceptualization: J.J.H-M. and B.R-M; Data curation: M.A. M-S. and I.R-M; Formal analysis: J.J.H-M. and H.P-S; Funding acquisition: B.R-M; Investigation: V.E.F-R., M-F-G, A.S., A.B-G., and M.D.F.; Methodology: J.J.H-M. and B.R-M.; Project administration: B.R-M.; Resources: B.R-M.; Software: H.P-S.; Supervision: B.R-M.; Validation: J.J.H-M. and B.R-M.; Visualization: M.A.M-S., V.E.F-R., M-F-G, M.D.F., and B.R-M. Writing – original draft: All authors. Writing – review and editing: All authors.

Conflicts of interest disclosure

All the authors declare to have no conflicts of interest relevant to this study.

Research registration unique identifying number (UIN)

Not applicable.

Guarantor

Juan José Hernández-Morante and Bruno Ramos-Molina accept full responsibility for the work and the conduct of the study, had access to the data, and controlled the decision to publish.

Provenance and peer review

Not commissioned, externally peer reviewed.

Data availability statement

Data are available under request to the corresponding author.

Declaration of AI use

The authors declare that AI-based tools were used solely for technical support during the preparation of this manuscript. Specifically, AI tools were employed to assist in debugging and refining R and Python codes, as well as to improve the grammar, spelling, and clarity of the written text. AI tools were not used for data generation, statistical analysis, interpretation of results, or decision-making. All analyses, interpretations, and conclusions are the sole responsibility of the authors.

Acknowledgements

We would like to thank Carlos Martínez and Antonio Banegas for their assistance with the machine learning analysis. We also thank Dr. Cristina Reche for her assistance with the final version of this paper. We are thankful to Microomics Systems S.L for their work with the metagenome analysis.

References

- [1] Boustani P, Sheidaei A, Mokhber S, *et al.* Assessment of weight change patterns following Roux en Y gastric bypass, one anastomosis gastric bypass and sleeve gastrectomy using change-point analysis. *Sci Rep* 2024;14:1–10.
- [2] Mantziari S, Thomopoulos T, Abboretti F, *et al.* Long-term weight loss and metabolic benefit from Roux-en-Y gastric bypass in patients with superobesity. *BJS Open* 2022;6:zrac145.
- [3] Abbott S, Adams B, Agrawal S, *et al.* Roux-en-Y gastric bypass, adjustable gastric banding, or sleeve gastrectomy for severe obesity (By-Band-Sleeve): a multicentre, open label, three-group, randomised controlled trial. *Lancet Diabetes Endocrinol* 2025;13:410–26.
- [4] Chang AR, Grams ME, Navaneethan SD. Bariatric surgery and kidney-related outcomes. *Kidney Int Rep* 2017;2:261.
- [5] Zhang HW, Han XD, Liu WJ, *et al.* Is Roux-en-Y gastric bypass advantageous—surgical outcomes in obese patients with type-2 diabetes after gastric bypass versus sleeve gastrectomy, a matched retrospective study. *Ann Transl Med* 2020;8:372.
- [6] Tasabehji D, Saleh S, Mokadem M. Impact of bariatric surgery and endoscopic therapies on liver health in metabolic dysfunction-associated steatotic liver disease: a review. *J Clin Med* 2025;14:4012.
- [7] Verrastro O, Panunzi S, Castagneto-gissey L, *et al.* Bariatric–metabolic surgery versus lifestyle intervention plus best medical care in non-alcoholic steatohepatitis (BRAVES): a multicentre, open-label, randomised trial. *Lancet* 2023;401:1786–97.
- [8] Still CD, Wood GC, Chu X, *et al.* Clinical factors associated with weight loss outcomes after Roux-en-Y gastric bypass surgery. *Obesity (Silver Sp)* 2014;22:888.
- [9] Biter LU, 't hart JW, Noordman BJ, *et al.* Long-term effect of sleeve gastrectomy vs Roux-en-Y gastric bypass in people living with severe obesity: a phase III multicentre randomised controlled trial (SleeveBypass). *Lancet Reg Health Eur* 2024;38:100836.
- [10] Aasbrenn M, Svendstrup M, Schnurr TM, *et al.* Genetic markers of abdominal obesity and weight loss after gastric bypass surgery. *PLoS One* 2021;16:e0252525.
- [11] Steffen KJ, Sorgen AA, Fodor AA, *et al.* Early changes in the gut microbiota are associated with weight outcomes over 2 years following metabolic and bariatric surgery. *Obesity* 2024;32:1985–97.
- [12] Juárez-fernández M, Román-sagüillo S, Porras D, *et al.* Long-term effects of bariatric surgery on gut microbiota composition and faecal metabolome related to obesity remission. *Nutrients* 2021;13:2519.
- [13] Palmas V, Pisanu S, Madau V, *et al.* Gut microbiota markers associated with obesity and overweight in Italian adults. *Sci Rep* 2021;11:1–14.
- [14] Georgiou K, Belev NA, Koutouratsas T, *et al.* Gut microbiome: linking together obesity, bariatric surgery and associated clinical outcomes under a single focus. *World J Gastrointest Pathophysiol* 2022;13:59.
- [15] Herz CT, Kulterer OC, Prager M, *et al.* Bariatric surgery promotes recruitment of brown fat linked to alterations in the gut microbiota. *Eur J Endocrinol* 2025;192:603–13.
- [16] Martínez-montoro JL, Kuchay MS, Balaguer-Román A, *et al.* Gut microbiota and related metabolites in the pathogenesis of nonalcoholic steatohepatitis and its resolution after bariatric surgery. *Obes Rev* 2022;23. doi:10.1111/OBR.13367
- [17] Gutiérrez-Repiso C, Garrido-sánchez L, Alcaide-torres J, *et al.* Predictive role of gut microbiota in weight loss achievement after bariatric surgery. *J Am Coll Surg* 2022;234:861–71.
- [18] Martínez-Sánchez MA, Balaguer-román A, Fernández-Ruiz VE, *et al.* Plasma short-chain fatty acid changes after bariatric surgery in patients with severe obesity. *Surg Obesity Related Dis* 2023;19:727–34.
- [19] Sutter T, Roth JA, Chin-cheong K, *et al.* A comparison of general and disease-specific machine learning models for the prediction of unplanned hospital readmissions. *J Am Med Inform Assoc* 2021;28:868–73.
- [20] Salazar N, Ponce-alonso M, Garriga M, *et al.* Fecal metabolome and bacterial composition in severe obesity: impact of diet and bariatric surgery. *Gut Microbes* 2022;14:2106102.
- [21] Gutiérrez-repiso C, Moreno-indias I, De hollanda A, *et al.* Gut microbiota specific signatures are related to the successful rate of bariatric surgery. *Am J Transl Res* 2019;11:942.
- [22] Casas domínguez M, Herrera montano I, López Gómez JJ, *et al.* Predicting weight loss success after gastric sleeve surgery: a machine learning-based approach. *Nutrients* 2025;17:1391.
- [23] Huang X, Marques-silva J. On the failings of Shapley values for explainability. *Int J Approx Reason* 2024;171:109112.
- [24] Pantelis AG. Metabolomics in bariatric and metabolic surgery research and the potential of deep learning in bridging the gap. *Metabolites* 2022;12:458.
- [25] Eisenberg D, Shikora SA, Aarts E, *et al.* 2022 American Society of Metabolic and Bariatric Surgery (ASMBS) and International Federation for the Surgery of Obesity and Metabolic Disorders (IFSO) indications for metabolic and bariatric surgery. *Obes Surg* 2022;33:3.
- [26] Martínez-sánchez MA, Ros-madrid I, Fernández-ruiz VE, *et al.* Mid-term changes in quality of life and nutritional habits following gastric bypass: a 24-month follow-up study. *Nutrients* 2026;18:288.
- [27] Agha RA, Mathew G, Rashid R, *et al.* Transparency in the reporting of Artificial intelligence –the TITAN guideline. *Prem J Sci* 2025;10:100082.
- [28] Sabench pereferer F, Molina lópez A, Acín gándara D, *et al.* Analysis of mid-term weight loss after roux-en-y gastric bypass and sleeve gastrectomy: proposed percentile charts for total weight loss percentages to be used in daily clinical practice. *Cirugía Española (Engl Ed)* 2020; 98:72–78.
- [29] Pruesse E, Quast C, Knittel K, *et al.* SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 2007;35:7188–96.
- [30] Debelius JW, Robeson M, Hugerth LW, *et al.* A comparison of approaches to scaffolding multiple regions along the 16S rRNA gene for improved resolution. *bioRxiv* 2021:2021–23. doi:10.1101/2021.03.23.436606
- [31] Banegas-luna AJ, Pérez-sánchez H. SIBILA: automated machine-learning-based development of interpretable machine-learning models on high-performance computing platforms. *AI* 2024;5:2353–74.
- [32] Chen M, Tang TC, Chen YY, *et al.* A meta-analysis of changes in gut microbiota structure following bariatric surgery. *Int J Surg* 2025;111: 8480–92.
- [33] Morán-ramos S, Soriano-cortés R, Soto-fuentes V, *et al.* Role of pre-surgical gut microbial diversity in Roux-en-Y Gastric bypass weight-loss response: a cohort study. *Lifestyle Genom* 2024;17:12–21.
- [34] Pinart M, Dötsch A, Schlicht K, *et al.* Gut microbiome composition in obese and non-obese persons: a systematic review and meta-analysis. *Nutrients* 2022;14:12.
- [35] Grahnmemo L, Nethander M, Coward E, *et al.* Cross-sectional associations between the gut microbe *Ruminococcus gnavus* and features of the metabolic syndrome. *Lancet Diabetes Endocrinol* 2022;10:481–83.
- [36] Zhou H, Tang L, Fenton KA, *et al.* Exploring and evaluating microbiome resilience in the gut. *FEMS Microbiol Ecol* 2025;101:46.
- [37] Visconti A, Le roy CI, Rosa F, *et al.* Interplay between the human gut microbiome and host metabolism. *Nat Commun* 2019;10:4505–.
- [38] Davies NK, O'sullivan JM, Plank LD, *et al.* Altered gut microbiome after bariatric surgery and its association with metabolic benefits: a systematic review. *Surg Obesity Related Dis* 2019;15:656–65.
- [39] Liu R, Hong J, Xu X, *et al.* Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nature Med* 2017;23:859–68.

- [40] Gámez-valdez JS, Corona-cervantes K, Sánchez-salguero ES, *et al.* Analysis of human colostrum reveals differential co-occurrence networks of metabolites, microbiota and cytokines in maternal obesity. *Food Funct* 2025;16:5900–16.
- [41] Peters BA, Shapiro JA, Church TR, *et al.* A taxonomic signature of obesity in a large study of American adults. *Sci Rep* 2018;8:9749–.
- [42] Castaner O, Goday A, Park YM, *et al.* The gut microbiome profile in obesity: a systematic review. *Int J Endocrinol* 2018;2018:4095789.
- [43] Masoumi M, Jafarzadeh A, Hadavi H, *et al.* Human gut microbiome before and after bariatric surgery in obese patients with and without type 2 diabetes. *Obes Surg* 2024;34:2835–43.
- [44] Luijten JCHBM, Vugts G, Nieuwenhuijzen GAP, *et al.* The importance of the microbiome in bariatric surgery: a systematic review. *Obes Surg* 2019;29:2338–49.
- [45] Pedroso I, Kumbhare SV, Johnson S, *et al.* Microbiome and genetic predictors of weight loss 12 months postsleeve gastrectomy: insights from a pilot retrospective cohort study. *Surg Obesity Related Dis* 2025;21:1035–41.
- [46] Zhang L, Cheng X, Xia L, *et al.* Analysis of 16s rRNA gene sequencing in feces: the impact of bariatric surgery on the gut microbiota in patients with obesity. *Obes Surg* 2024;34:1185–95.
- [47] Ben Izhak M, Eshel A, Cohen R, *et al.* Projection of gut microbiome pre- and post-bariatric surgery to predict surgery outcome. *mSystems* 2021;6. doi:10.1128/MSYSTEMS.01367-20
- [48] Fouladi F, Carroll IM, Sharpton TJ, *et al.* A microbial signature following bariatric surgery is robustly consistent across multiple cohorts. *Gut Microbes* 2021;13. doi:10.1080/19490976.2021.1930872
- [49] Yadav J, Liang T, Qin T, *et al.* Gut microbiome modified by bariatric surgery improves insulin sensitivity and correlates with increased brown fat activity and energy expenditure. *Cell Rep Med* 2023;4:101051.
- [50] Gopalakrishnan V, Kumar C, Robertsen I, *et al.* A multi-omics microbiome signature is associated with the benefits of gastric bypass surgery and is differentiated from diet induced weight loss through 2 years of follow-up. *Mucosal Immunol* 2025;18:825–35.
- [51] Coimbra VOR, Crovesy L, Ribeiro-alves M, *et al.* Gut microbiota profile in adults undergoing bariatric surgery: a systematic review. *Nutrients* 2022;14:4979.
- [52] Ngowi EE, Wang YZ, Khattak S, *et al.* Impact of the factors shaping gut microbiota on obesity. *J Appl Microbiol* 2021;131:2131–47.
- [53] Chen VP, Gao Y, Geng L, *et al.* Butyrylcholinesterase deficiency promotes adipose tissue growth and hepatic lipid accumulation in male mice on high-fat diet. *Endocrinology* 2016;157:3086–95.
- [54] Gomez-arango LF, Barrett HL, McIntyre HD, *et al.* Connections between the gut microbiome and metabolic hormones in early pregnancy in overweight and obese women. *Diabetes* 2016;65:2214–23.
- [55] Kim MH, Yun KE, Kim J, *et al.* Gut microbiota and metabolic health among overweight and obese individuals. *Sci Rep* 2020;10:19417–.