Computational modeling enables individual assessment of postprandial glucose and insulin responses after bariatric surgery

Onur Poyraz[1*] **Sini Heinonen**[2,3*] **S. T. John**[1] **Tuure Saarinen**[4] **Anne Juuti**[4]
Pekka Marttinen[1†] **Kirsi H. Pietiläinen**[2,5†]
[1]Department of Computer Science, Aalto University
[2]Obesity Research Unit, Research Programs Unit, Clinical and Molecular Metabolism, University of Helsinki
[3]Department of Internal Medicine, Helsinki University Hospital
[4]Department of Gastrointestinal Surgery, Abdominal Center, Helsinki University Hospital and University of Helsinki
[5]HealthyWeightHub, Endocrinology, Abdominal Center, Helsinki University Hospital and University of Helsinki
onur.poyraz@aalto.fi
sini.heinonen@helsinki.fi

Abstract

Bariatric surgery enhances glucose metabolism, yet the detailed postprandial joint glucose and insulin responses, variability in individual outcomes, and differences in surgical approaches remain poorly understood. To address this, we used hierarchical multi-output Gaussian process (HMOGP) regression to reveal clinically relevant patterns between persons undergoing two types of bariatric surgery by modeling the individual postprandial glucose and insulin responses and estimating the average response curves from individual data. 44 participants with obesity underwent either Roux-en-Y gastric bypass (RYGB; n=24) or One-Anastomosis gastric bypass (OAGB; n=20) surgery. The participants were followed up at the 6th and 12th months after the operation, during which they underwent an oral glucose tolerance test (OGTT) and a mixed meal test (MMT). A marked reduction in glycemia, an earlier glucose peak, and an increase and sharpening in the postprandial glucose and insulin responses were evident in both metabolic tests post-operation. MMT resulted in higher postprandial glucose and insulin peaks compared with OGTT. Higher glucose and insulin responses were observed after RYGB compared with OAGB, suggesting differences between the procedures that may influence the clinical practice. Computational modeling with HMOGP regression can thus be used to, in detail, predict the combined responses of patient cohorts to ingested glucose or a mixed meal and help in assessing individual metabolic improvement after weight loss. This can lead to new knowledge in personalized metabolic interventions.

1 Introduction

The prevalence of obesity has increased enormously over the last decades. Over 1.8 billion people lived with overweight or obesity in 2017 (WHO factsheet). Obesity is associated with various diseases and metabolic problems such as type 2 diabetes (T2DM), dyslipidemia, cardiovascular diseases, and cancer (Zimmet et al., 2005; Guh et al., 2009).

Treating obesity is extremely difficult. While diet-induced weight loss strategies lead to modest results, bariatric surgery is the most effective treatment of obesity, leading to improved glucose tolerance already within days (Adams et al., 2012; Sjöström, 2013). Roux-en-Y gastric bypass (RYGB) has been the gold standard in bariatric surgery. One-Anastomosis gastric bypass (OAGB) emerged as a potentially more beneficial procedure for the resolution of T2DM (Almalki et al., 2018; Magouliotis et al., 2018), but recent randomized controlled trials have challenged this (Disse et al., 2014; Robert

[*]These authors contributed equally to this work.
[†]These authors contributed equally to this work

Preprint. Under review.

et al., 2019; Heinonen et al., 2023). Additionally, individual metabolic phenotypes of participants undergoing bariatric surgery, such as differential glucose response between sexes (Mauvais-Jarvis, 2018), have resulted in the need for understanding and modeling in detail the glycemic responses in different patient cohorts before and after surgery.

Conventional methods for assessing glucose responses after the oral glucose tolerance test (OGTT) and the mixed meal test (MMT) have relied on averaging the means of patient responses for each individual separately. Previously used methods in assessing the results of these tests therefore yield a robust estimation of blood glucose regulation but lack detailed characteristics of the overall blood glucose/insulin response at the population level and neglect individual differences. Additionally, OGTT is sparsely sampled and individual estimates may be missing from the available data. Artificial intelligence techniques to better understand blood glucose regulation have emerged in recent years (Li et al., 2020). For example, the Eindhoven Diabetes Education Simulator (E-DES) model has been successfully used for OGTT to develop personalized models of glucose and insulin sensitivity and to explore the heterogeneity in the responses (Erdős et al., 2021). In another study, methods quantifying glucose fluxes between tissues using tracers have been applied to create predictive models of glucose-insulin interactions in MMT analyses (Dalla Man et al., 2007). Zeevi et al. (2015) developed a machine-learning algorithm integrating blood parameters, dietary habits, anthropometrics, physical activity, and gut microbiota to estimate the glucose and insulin responses accurately. A Gaussian process has been previously used in modeling OGTT responses in participants with and without cystic fibrosis to provide measures of beta-cell function with quantified uncertainty (Garrish et al., 2023).

In this study, we developed hierarchical multi-output Gaussian process (HMOGP) regression, which is the combination of hierarchical Gaussian process (HGP) and multi-output Gaussian process (MOGP). The hierarchical Gaussian process (HGP) is beneficial when part of the data reveals different characteristics, such as the considerable inter-individual variability usually observed in postprandial glucose and insulin responses (Park and Choi, 2010). On the other hand, the multi-output Gaussian process (MOGP) is able to deal with multiple correlated outputs (such as glucose and insulin dynamics) and to provide more accurate predictions than simply modeling these outputs separately (Álvarez et al., 2010; Liu et al., 2018; Moreno-Muñoz et al., 2018). Similar techniques to HMOGP have been successfully applied before in different domains such as anomaly detection (Cho et al., 2019), multi-task learning (Li and Chen, 2018), and mobile cellular network analysis (Cuesta Ramírez, 2015), performing better than MOGP and GP. We expanded these techniques to the medical domain in OGTT and MMT responses with necessary adaptations. HMOGP allows us (1) to model the insulin and glucose response together via multi-output (2) while considering the individual level differences through a hierarchical structure. It also (3) quantifies uncertainty in predictions, making it suitable for datasets with noise or missing values, and (4) can model non-linearities and complex temporal dynamics in postprandial responses, providing more accurate predictions than simpler models (Williams and Rasmussen, 2006). Modeling both OGTT and MMT data in detail with HMOGP regression allows for a more detailed assessment of postprandial glucose and insulin responses. We evaluated model outcomes to better understand the postprandial glucose and insulin levels after an OGTT and an MMT, considering before and after different types of bariatric surgery and between sexes or participants with or without T2DM, providing a thorough analysis of postprandial response with an advanced technique.

2 Methods

2.1 Subjects

We included participants with obesity (n=44, aged 46.6 years with 14 men and 30 women, 26 participants with obesity, 18 participants with obesity and T2DM, matched for age and diabetes status between the operation types). Participants were examined before and at 6th and 12th months after bariatric surgery with either Roux-en-Y gastric bypass (RYGB; n=24, 8 men, 16 women, n=9 with T2DM) or One-Anastomosis gastric bypass (OAGB; n=20, 6 men, 14 women, n=9 with T2DM). 33% of men vs. 43% of women had T2DM. The subjects were recruited from Helsinki University Hospital through the Department of Gastrointestinal Surgery. The full randomized clinical trial is registered at clinicaltrials.gov with no. NCT02882685 and the randomization and selection process of the participants and study dropouts are described in detail in Saarinen et al. (2019) and Heinonen et al. (2023). All participants from Helsinki University Hospital with OGTT data available were included in the current study. No data or participants were removed from the analysis. Written informed consent was obtained from all participants. The study protocol was designed and performed according to the principles of the Helsinki Declaration and approved by the Ethical Committee of the Helsinki University Central Hospital.

2.2 Bariatric surgery

In RYGB, the gastric pouch was created with one horizontal 45 mm and two vertical 60 mm staplers. The length of the biliary limb was 80 cm, and that of the alimentary limb was 130 cm. In OAGB, a tubular gastric pouch was

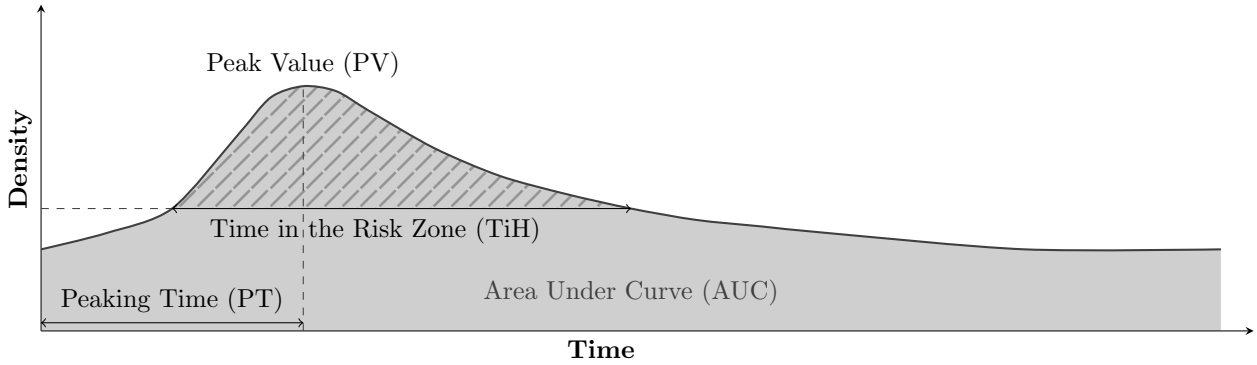

Figure 1 | Definition of the metrics evaluated in the experiments. Peaking Time (PT) and Time in the Risk Zone (TiH; Hyperglycemia for glucose > 7.0mmol/l and Hyperinsulinemia for insulin > 50mU/l) represented in minutes, and Peak Value (PV) and time-normalized Area Under Curve (AUC) represented in mmol/l for glucose and mU/l for insulin.

created using 60 mm staplers along a 38Fr bougie starting at the crow's foot with a horizontal 45 mm stapler and the omega loop being 210 cm long. The length of the bypasses was standardized between the procedures to allow for equal comparison. A 210 cm biliopancreatic limb in OAGB and 80 cm biliopancreatic and 130 cm alimentary limbs in RYGB were chosen to obtain equally long bypassed intestines in both groups (Saarinen et al., 2019).

2.3 Clinical examinations and body composition

Weight and height were measured after an overnight fast in light clothing. Whole body composition was measured by Dual-energy X-ray absorptiometry (DEXA) using a Lunar Prodigy whole-body scanner (GE Medical Systems, Madison, WI).

2.4 Analytical blood samples

Fasting laboratory tests, including Hba1c and plasma lipids, were performed as described by Heinonen et al. (2015).

2.5 Mixed meal test (MMT) and oral glucose tolerance test (OGTT)

Glucose metabolism was measured after an overnight fast with a 3-hour oral glucose tolerance test (OGTT, 75 g of glucose), with time points 0 (before the ingestion of the glucose drink) and post-glucose samples at 30, 60, 120, and 180 min. We additionally performed a 6-hour Mixed Meal Test (MMT) on a separate day, where a fasting blood sample was collected before ingesting the liquid meal of 2620 kJ (627 kcal) with a balanced distribution of fat (24 g), carbohydrates (76 g) and protein (24 g) (Resource® 2.5 Compact, Nestle Health Science), with post-meal samples at 15, 30, 60, 120, 180, 240 and 360 min for measuring glucose and insulin.

2.6 Data features

Each patient underwent both OGTT and MMT tests across three different visits: once preoperatively and then at the 6th and 12th months postoperatively. Six participants missed the baseline OGTT visit, while 10 and 6 missed the 6th and 12th month visits, respectively. In MMT, 4 and 5 participants missed the 6th and 12th month visits, respectively. Additionally, 3% of the overall data was missing due to vomiting or nausea, which interrupted the test.

2.7 Statistical method

Blood glucose and insulin evolve as a function of time. Here, we model them using a Gaussian process, which is a continuous stochastic process and can be considered a probability distribution over functions (MacKay et al., 1998; Williams and Rasmussen, 2006). The standard Gaussian process regression is defined as:

$$\boldsymbol{y} \sim \mathcal{N}\left(f(t), \sigma I\right), \tag{1}$$

where $\boldsymbol{y} \in \mathbb{R}^T$ are the observation points, $t \in \{1, \ldots, T\}$ the input locations, and σ the measurement error. The function $f(t)$ is defined as the Gaussian process $f(t) \sim \mathcal{GP}(0, \mathbf{K})$ with kernel matrix $\mathbf{K}$. However, applying this approach directly to our problem is not straightforward. First, we know that glucose and insulin responses are strongly

correlated, and applying Gaussian process regression on them separately would miss this correlation. Therefore, we propose to use Multi-Output Gaussian Process regression (Álvarez et al., 2010; Liu et al., 2018; Moreno-Muñoz et al., 2018), which is able to calculate the cross-covariance between functions and is the proper choice for modeling insulin and glucose together. If we define $\boldsymbol{f} = [f_{\text{glucose}}, f_{\text{insulin}}]$:

$$f_{\text{glucose}}(t) = \mathbf{L}_{11}\hat{f}_{\text{glucose}}(t) \tag{2}$$

$$f_{\text{insulin}}(t) = \mathbf{L}_{21}\hat{f}_{\text{glucose}}(t) + \mathbf{L}_{22}\hat{f}_{\text{insulin}}(t) \tag{3}$$

where $\hat{f}_{\text{glucose}}, \hat{f}_{\text{insulin}} \in \mathbb{R}^T$ are independent GPs and $\mathbf{L}$ is the Cholesky factor matrix of the corresponding correlation matrix $\mathbf{C}$, then $\boldsymbol{f} \in \mathbb{R}^{T\times D}$, where D is the output dimension (glucose, insulin), has the multi-output GP distribution. However, the challenge of the data is not limited to the output correlation. The second issue is that individuals have slightly different blood glucose and insulin responses. Hierarchical Gaussian process regression modeling can solve this issue by learning the population average response curve and individual deviations from the average response separately (Park and Choi, 2010). Let $\boldsymbol{f}_n(t)$ denote the response curve for a single patient, $\boldsymbol{f}^{\text{population}}(t)$ the average response curve across all participants, and $\boldsymbol{f}_n^{\text{patient}}(t)$ the deviation of the patient from the average response. Then, hierarchical GP can be represented as:

$$\boldsymbol{f}_n(t) = \boldsymbol{f}^{\text{population}}(t) + \boldsymbol{f}_n^{\text{patient}}(t) \tag{4}$$

where $\boldsymbol{f}^{\text{population}}, \boldsymbol{f}_n^{\text{patient}} \in \mathbb{R}^{T\times D}$ have multi-output Gaussian process distributions. We apply the same hierarchical approach to model the constant terms for fasting glucose and insulin, as follows:

$$\boldsymbol{\mu}_n = \boldsymbol{\mu}^{\text{population}} + \boldsymbol{\mu}_n^{\text{patient}} \tag{5}$$

where $\boldsymbol{\mu}_n$ is fasting glucose and insulin for a single patient, $\boldsymbol{\mu}^{\text{population}}$ corresponds to the normally distributed population average of fasting glucose and insulin, and $\boldsymbol{\mu}_n^{\text{patient}}$ denotes the normally distributed individual deviation from the population average.

Overall, we modeled the data using the Hierarchical Multi-Output Gaussian process (HMOGP), where hierarchy allows us to consider the deviations of individual responses, and multi-output models the correlation of insulin and glucose response. Let the whole data be denoted by $\boldsymbol{y} \in \mathbb{R}^{N\times T\times D}$, where N is the total number of participants, T is the number of time points and $D = 2$ refers to the two possible outputs (glucose, insulin). If we index patient by n, time by t, and output by d, the blood glucose and insulin level $\boldsymbol{y}_{n,t}$ of patient n at time t can be represented as

$$\boldsymbol{y}_{n,t} \sim \mathcal{N}\left(\boldsymbol{f}_n(t), \Sigma\right), \tag{6}$$

where Σ is the diagonal measurement error matrix, and the latent function $\boldsymbol{f}_n$ is formulated as

$$\boldsymbol{f}_n(t) = \boldsymbol{\mu}^{\text{population}} + \boldsymbol{\mu}_n^{\text{patient}} + \boldsymbol{f}^{\text{population}}(t) + \boldsymbol{f}_n^{\text{patient}}(t). \tag{7}$$

We applied this model separately to the different visits (Baseline, 6th month, 12th month) and tests (OGTT, MMT). The only shared parameter across these models was the measurement error. For comparisons by operation, sex, and T2DM, we divided the dataset accordingly and modeled them separately. Further details of the model parameters and priors, along with a link to the Stan code of the model, can be found in Appendix A.

Stan software (Carpenter et al., 2017) was used to draw samples from the posterior distributions using the Markov Chain Monte Carlo inference. We used posterior samples to calculate Bayesian credible intervals (CI) and p-values as described in Gelman et al. (2013). Furthermore, we evaluated four metrics to investigate the characteristics of the postprandial glucose and insulin responses: peak value (PV), time-normalized area under the curve (AUC), peaking time (PT), and time in the risk zone (TiH; Hyperglycemia for glucose $>$ 7.0mmol/l and Hyperinsulinemia for insulin $>$ 50mU/l). These metrics are illustrated in Figure 1. In the clinical analyses, non-normally distributed variables were log10 transformed before the parametric analyses. Differences at baseline between the groups were analyzed by Student's T-tests and between visits with generalized linear mixed modeling adjusting for sex, T2DM, and operation type (Stata Statistical software ver. 7.0). These results are presented as means and 95% confidence intervals (CI).

3 Results

Throughout the manuscript, we investigated the glucose and insulin responses of all participants following an oral glucose tolerance test (OGTT) and a mixed meal test (MMT) over the entire dataset and in different subgroups divided by operation, sex, and T2DM using hierarchical multi-output Gaussian process (HMOGP) regression, which is suitable for longitudinal data. In MMT, 76g of glucose was combined with protein and fats, while OGTT had only 75g of glucose. Figure B.1 and Figure B.2 show HMOGP predictions by individuals, which validates the model fit for different persons. Note that Figure B.1 and Figure B.2 show only the participants with the complete set of observations while the model was fitted using all the participants.

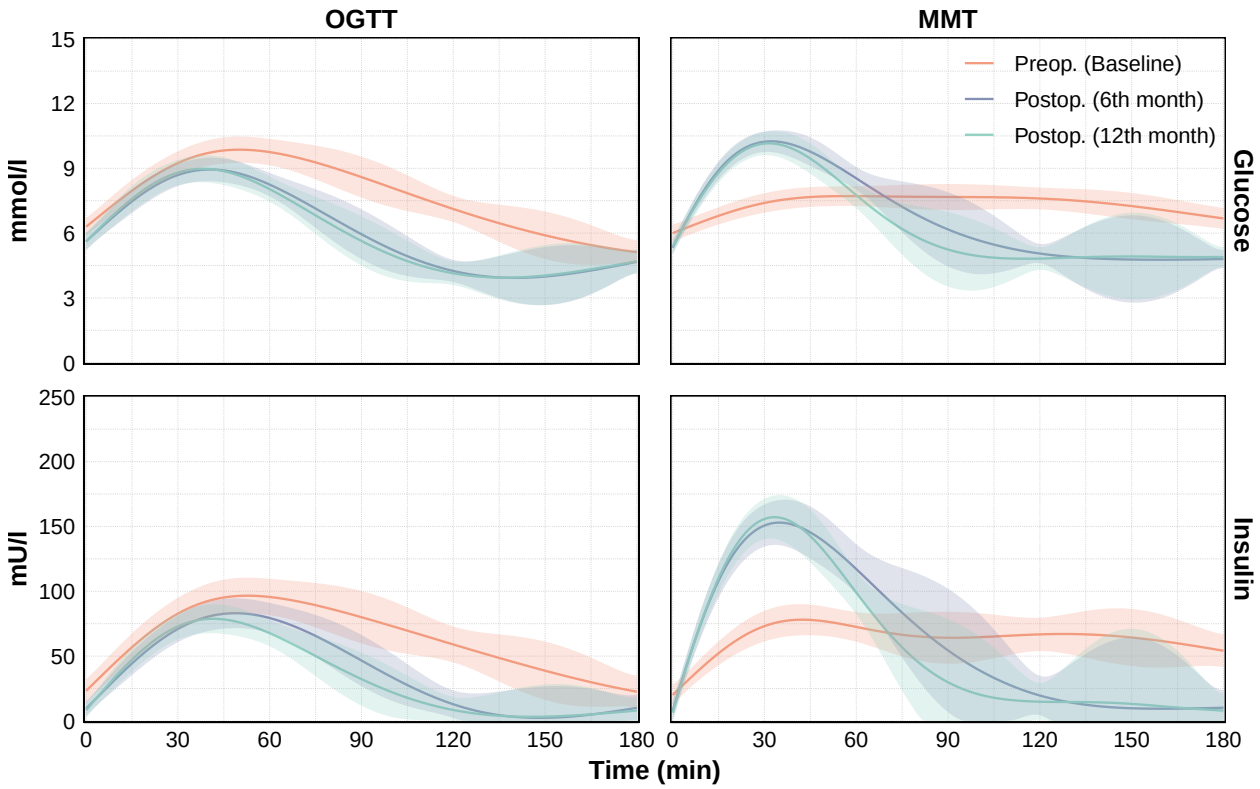

Figure 2 | The comparison of model predicted average glucose and insulin responses in different visits (Baseline, 6th month, and 12th month) within OGTT and MMT tests. Shaded areas represent 95% confidence intervals (CI).

3.1 Glucose and insulin responses sharpen post-surgery in OGTT and MMT

From baseline to the 12th month of post-surgery, the glucose-insulin response improved in both tests (Figure 2). During OGTT, AUCs of glucose and insulin concentrations significantly decreased from baseline to 6th and 12th months (Table F.3). Similarly, PVs of both glucose and insulin were reduced significantly from baseline to the 12th month (Table F.2), and PT at the 12th month occurred 12 minutes earlier for glucose and 11 minutes earlier for insulin compared to baseline (Table F.1). TiH for glucose was reduced by 53 min during weight loss, already bringing about large metabolic benefits (Table F.4). Overall, OGTT demonstrated a sharpened and significantly earlier glucose and insulin response with lower peak values following bariatric surgery (Figure 2).

In MMT, AUCs of both glucose and insulin showed a significant reduction from baseline to the 12th month (Table F.3). Unlike in OGTT, PVs of glucose and insulin significantly increased over 12 months (Table F.2). PT occurred 48 minutes and 11 minutes earlier for glucose and insulin, respectively, from baseline to the 12th month (Table F.1). Additionally, there was a total reduction of 83 minutes in TiH for glucose from baseline to the 12th month (Table F.4). These findings indicate a sharper and more pronounced glucose-insulin response, with significantly higher peak values and earlier peak times after bariatric surgery in MMT (Figure 2).

3.2 Higher postprandial glucose-insulin response in MMT vs. OGTT after bariatric surgery

OGTT and MMT evaluate participants' metabolic status through differential ingestion of macro-nutrients between the tests. To elucidate the differences between the tests before and after bariatric surgery, we next analyzed the average postprandial response in glucose and insulin, comparing MMT and OGTT by including similar follow-up times of 180 minutes in both tests (Figure 3). PVs of glucose and insulin were higher after OGTT than MMT at baseline, but other metrics did not differ significantly at baseline (Appendix F). Post-operatively at the 6th and 12th months, glucose and insulin responses (AUC, PV, and PT) were all significantly enhanced in MMT compared with OGTT (Appendix F). TiH for glucose did not differ between the tests during weight loss. While both tests are performed in the same individuals and on following days, these differences could be attributed to differential absorption of glucose in combination with other macro-nutrients in MMT vs. OGTT post-operatively and to postprandially enhanced insulin stimulus originating from proteins and fats included in MMT.

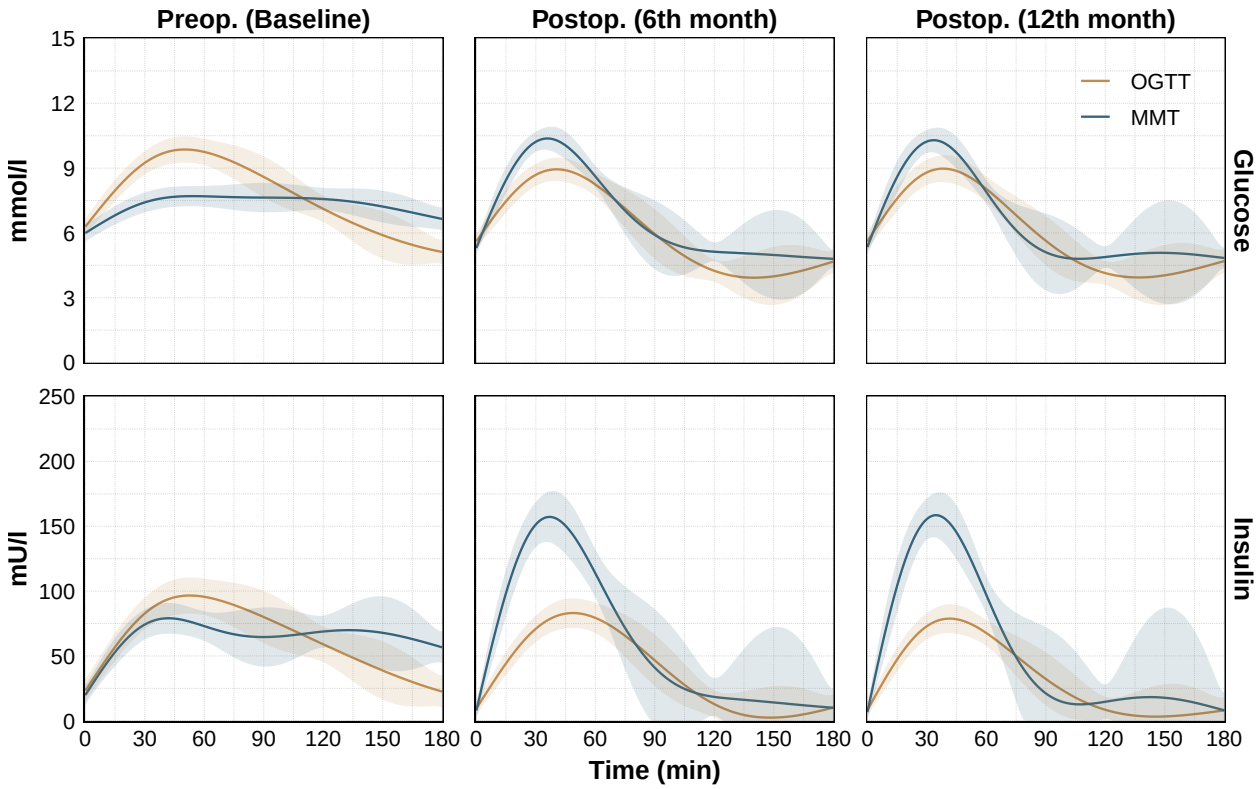

Figure 3 | The comparison of model predicted average glucose and insulin responses by OGTT and MMT test in different visits. Shaded areas represent 95% confidence intervals (CI).

3.3 Glucose-insulin response is higher and longer after RYGB than OAGB

To shed light on the potential differential responses between two types of bariatric surgery, we modeled RYGB (n=24) and OAGB (n=20) groups separately. Figure 4 and Figure 5 show population comparison for OGTT and MMT tests, respectively. In OGTT, the RYGB and OAGB groups had comparable glucose and insulin responses at baseline. The RYGB group showed a trend towards a larger AUC of glucose compared to the OAGB group (p=0.09) at the 12th month and had a significantly larger AUC of insulin at the 6th month (Table F.3). PV of glucose for the RYGB group was higher than the OAGB group at the 12th month (10.0 mmol/l vs. 7.8 mmol/l). For insulin, the RYGB group had a higher PV at the 6th month and a trend towards it at the 12th month (p=0.062, Table F.2). PTs showed no difference between operations (Table F.1)). TiH for glucose was longer in RYGB vs. OAGB group at the 12th month (70 min vs. 47 min), but this difference was nearly significant already at baseline (125 min vs. 105 min, p=0.088, Table F.4). During the entire weight-loss period, the mean TiH reduction of glucose was similar between RYGB and OAGB (55 min vs. 58 min).

In MMT, the RYGB group had similar PV of glucose but slightly higher PV of insulin than the OAGB group at baseline (Table F.2). During weight loss, the enhanced glucose and insulin responses in RYGB vs. OAGB were pronounced, with significantly higher PVs (Table F.2) and AUCs (Table F.3) of both glucose and insulin at 6th and 12th months. PTs of glucose and insulin were similar between the groups at all visits (Table F.1). Interestingly, the RYGB group experienced more TiH for glucose at the 6th month (83 min vs. 53 min) and a trend towards it at the 12th month (69 min vs. 54 min, p=0.054) compared with the OAGB group, despite similar baseline values between the groups (Table F.4). This may be partly due to the higher PV of glucose observed in RYGB. Over the 12 months, the reduction in TiH for glucose was 64 minutes for RYGB and 82 minutes for OAGB.

Overall, RYGB surgery induced a more pronounced glucose and insulin response in both tests, particularly in MMT, compared with OAGB, suggesting clinical differences between the procedures that can be taken into account in operation selection. Additionally, the pronounced PV and AUC in the RYGB group may drive the differences between MMT and OGTT in the postprandial response after surgery.

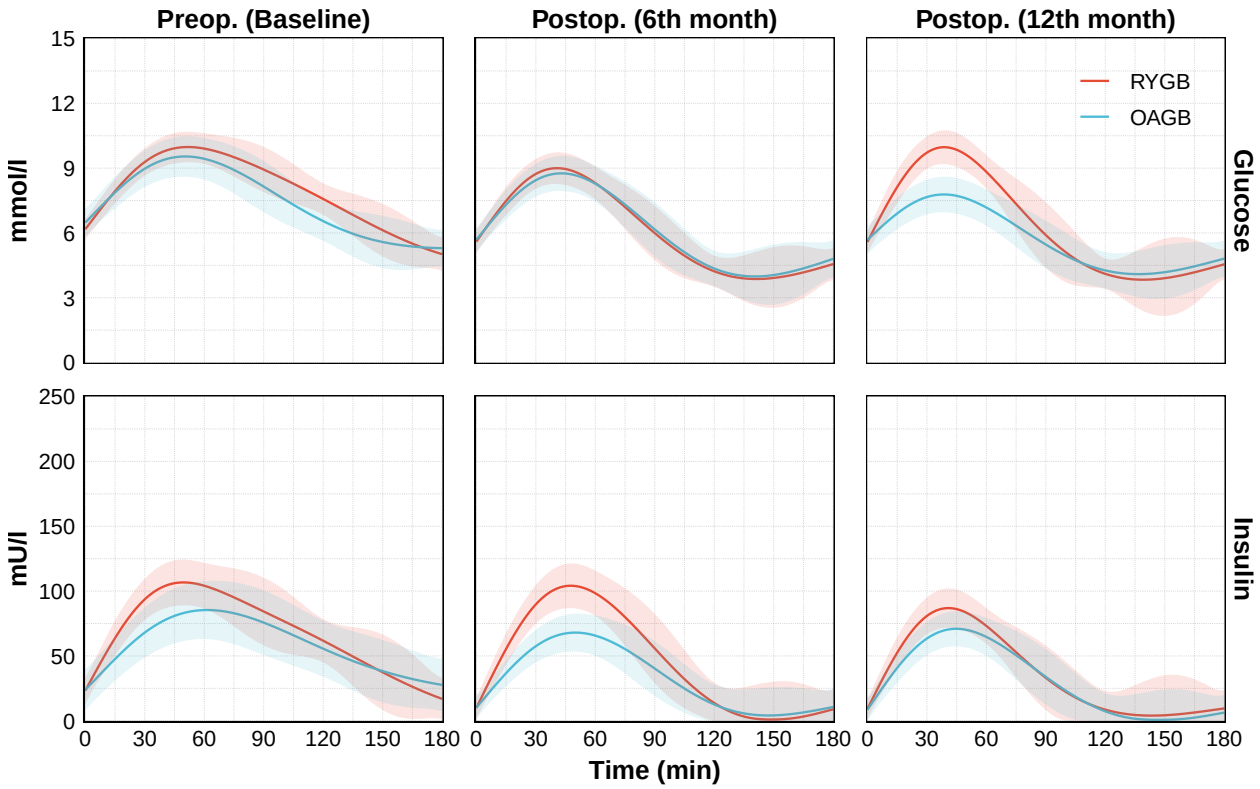

Figure 4 | The comparison of model predicted average glucose and insulin response of the participants with Roux-en-Y gastric bypass (RYGB) and One-Anastomosis gastric bypass (OAGB) operation in OGTT test. Shaded areas represent 95% confidence intervals (CI).

3.4 Higher postprandial glucose but not insulin response in men vs. women

Next, to account for other metabolic subgroups included in our cohort, we modeled women (n=30) and men (n=14) as different groups. Figure E.1 and Figure E.2 show population comparison for OGTT and MMT tests, respectively. In OGTT, men exhibited a higher AUC of insulin and higher PVs of glucose and insulin compared to women at baseline (Table F.3). Men had consistently higher postprandial PV of glucose response at all visits (Table F.2). At the 12th month, men had a higher AUC of glucose (Table F.3), later PTs for both glucose and insulin (Table F.1), and longer TiH for glucose with a smaller reduction in hyperglycemic time compared with women (31 min vs. 62 min) from baseline to 12th month (Table F.4).

In MMT, men and women showed similar baseline characteristics (Figure E.2). Men exhibited higher AUCs of glucose and insulin at the 6th month (Table F.3) along with higher PV of glucose at the 6th and 12th months (Table F.2) compared with women. At the 6th month, the PTs of glucose and insulin occurred later with men (Table F.1), and men also had longer TiH for glucose. However, the sexes had similar improvement in TiH during the weight loss (Table F.4).

3.5 Higher postprandial glucose but lower insulin response in T2DM vs. non-T2DM

We additionally evaluated the postprandial glucose and insulin responses between persons with T2DM (n=18) separately from those without (n=26) (Appendix D). As expected in OGTT, at all visits, persons with T2DM had higher PV and AUC, later PT and longer TiH for glucose, and later PT and lower PV for insulin (Appendix F). In MMT, PV and AUC of glucose were higher, and TiH for glucose was longer with T2DM at all visits. PV of insulin was lower than in persons without T2DM at the 12th month (Appendix F). Both groups improved in their glucose and insulin responses during weight loss.

3.6 Metabolic features in all participants improve during weight loss

At baseline, both RYGB and OAGB groups had similar metabolic features. Women had a higher fat percentage than men, and persons with T2DM had higher Hba1c and fasting plasma glucose than persons without T2DM. During the

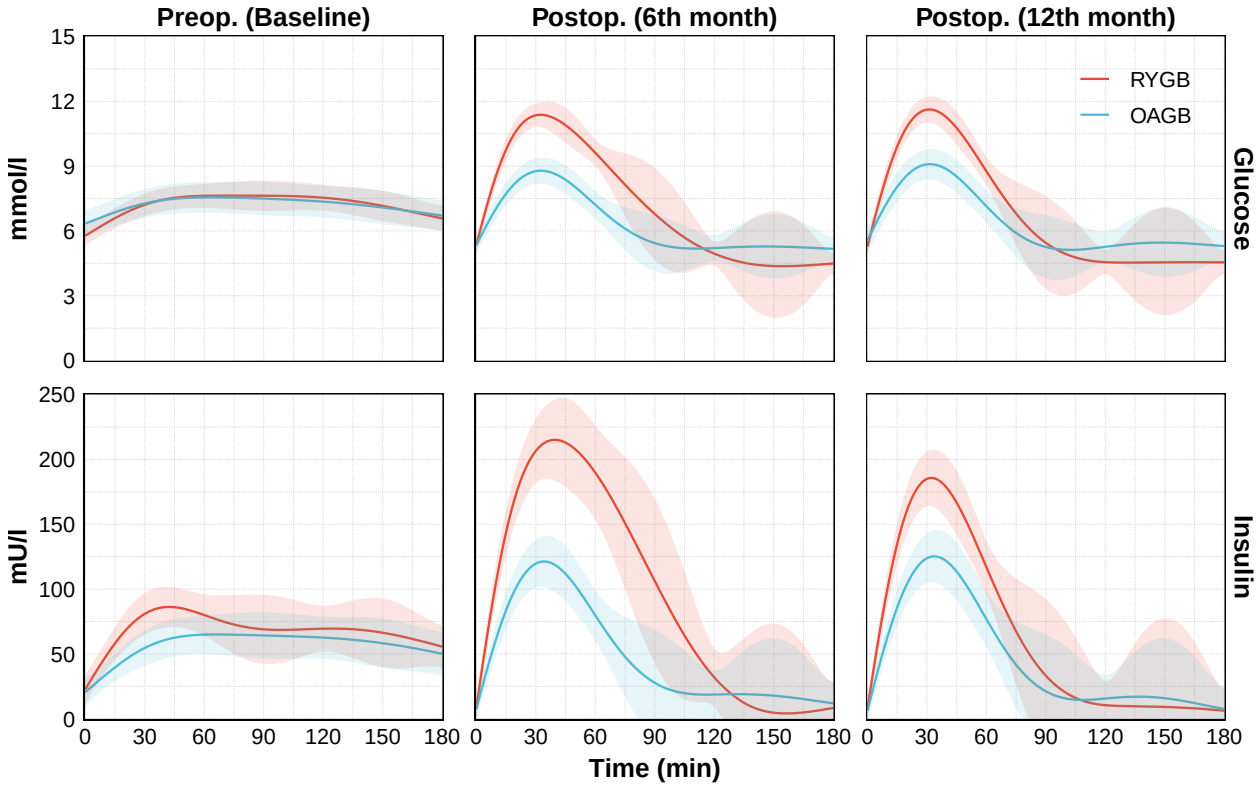

Figure 5 | The comparison of model predicted average glucose and insulin response of the participants with Roux-en-Y gastric bypass (RYGB) and One-Anastomosis gastric bypass (OAGB) operation in MMT test. Shaded areas represent 95% confidence intervals (CI).

weight loss, body weight, fat kilograms, fat percentage, fasting glucose, insulin, and lipid levels improved significantly in all groups (Table G.1). A slightly more prominent decrease in Hba1c in favor of OAGB vs. RYGB was observed, and HDL-cholesterol, fat kilograms, and fat percentage improved better in men vs. women. In persons with T2DM weight, Hba1c and fasting glucose improved more than in those without T2DM. (Table G.1).

3.7 HMOGP model outperforms the conventional trapezoidal data analysis

We additionally analysed the data by conventional methods of estimating PT, PV and calculating the AUC and TiH during the test with linear trapezoid method. The comparison with HMOGP showed that HMOGP consistently predicted AUC and TiH with improved uncertainty compared to the conventional data analysis. Additionally, HMOGP was able to give realistic PT and PV estimations compared to the linear trapezoid method, as these parameters only relied on the timing of the observations in fundamental statistical analysis. Appendix H shows the comparison of HMOGP predictions with trapezoidal data analysis.

4 Discussion

We applied hierarchical multi-output Gaussian process (HMOGP) regression to analyze glucose and insulin dynamics in OGTT and MMT, identifying distinct metabolic responses based on bariatric surgery type and sex. Our findings reveal sharper and earlier glucose-insulin peaks with faster restoration to baseline in both OGTT and MMT post-surgery, along with more pronounced postprandial glucose-insulin responses in MMT than OGTT and in RYGB compared to OAGB. Our findings underscore RYGB's role in enhanced glycemic control, while OAGB could be preferred in milder metabolic disturbances or when hypoglycemia may be a risk. RYGB's sharper glucose peaks may have vascular implications, favoring OAGB for participants with cardiovascular disease. Additionally, we observed sex-based metabolic differences with men exhibiting a more unhealthy glucose-insulin profile throughout weight loss and distinctions between individuals with and without T2DM. We thus present that the HMOGP approach can be used to model postprandial glucose-insulin responses in detail in different groups of participants for personalized treatment strategies, allowing new means of understanding and improving individual glycemic responses in metabolic conditions.

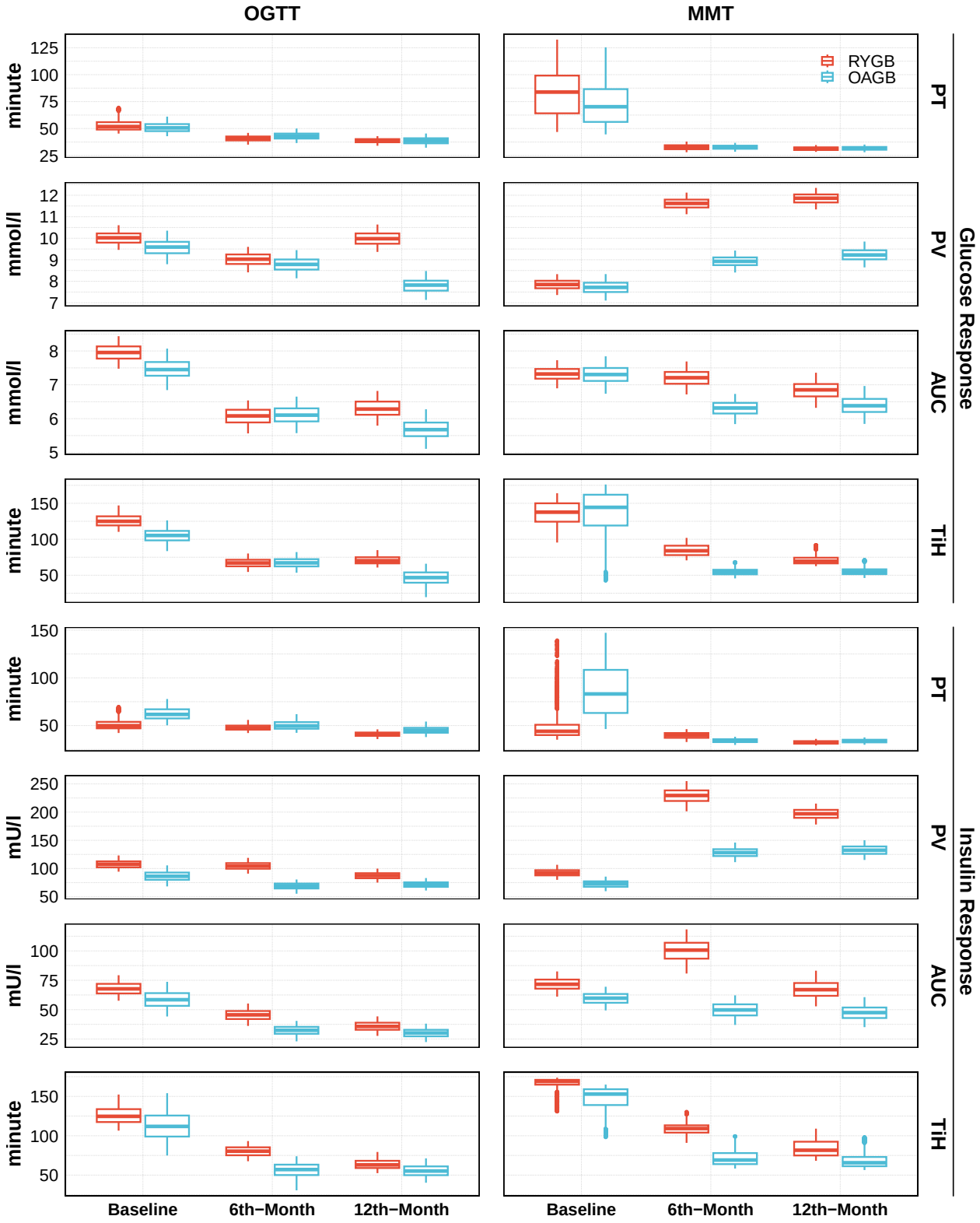

Figure 6 | The box-plot of each metric for operation comparison. See Appendix F for the 95 % CIs and p-values.

Our study proposes the hierarchical multi-output Gaussian process (HMOGP) regression machine learning model as a suitable tool for modeling postprandial glucose-insulin data in detail for both OGTT and MMT. Previous studies have used the standard Gaussian process for modeling OGTT responses in participants with and without cystic fibrosis (Garrish et al., 2023), but HMOGP has not been applied to OGTT or MMT in relation to obesity or weight loss. Traditional linear regression models or simple curve-fitting approaches assume independence between glucose and insulin and fail to capture individual variability. Non-hierarchical Gaussian processes do not handle individual variations as effectively as the hierarchical structure in HMOGP (Alvarez and Lawrence, 2011; Williams and Rasmussen, 2006). Additionally, fixed-effect models like ANOVA, or simpler mechanistic models, cannot accommodate individual-to-population hierarchy and multi-output dependencies. In conventional linear trapezoid calculations, the outcomes rely strongly on the exact measurement time points. Our HMOGP method will allow a more granular understanding of how a patient compares with the population in insulin and glucose responses after surgery. The additional advantages are suitability for the correlated structure of glucose and insulin dynamics, for datasets with noise or missing values (typical in OGTT and MMT), improved accuracy of the prediction, and the ability to handle complex, multivariate data.

Both OGTT and MMT assess the pancreatic beta-cell response, which includes activation of the gut-related insulin-stimulating hormones. While OGTT is the standard for evaluating glucose intolerance and T2DM, MMT provides a more comprehensive physiological stimulus, as glucose combined with protein and fats enhance insulin secretion (Cersosimo et al., 2014). The duration of the postprandial glucose and insulin peaks were significantly elevated and shortened after both metabolic tests, confirming enhanced insulin stimulus and improved glucose clearance from the circulation post-operatively. Our results align with previous studies reporting that RYGB leads to a rapid increase in both postprandial blood glucose and insulin levels (Ferrannini and Mingrone, 2009). These metabolic benefits likely stem from gut reconfiguration post-surgery, including altered nutrient absorption and enhanced stimulation of L-cells, leading to a 10-fold increase in insulin-stimulating GLP-1 secretion (Cummings et al., 2007), and an increase in other gut hormones (Sandoval and Patti, 2023). Gut mucosa hyperplasia post-operation promotes glucose absorption (Cavin et al., 2016) while improved pancreatic islet coordination may enhance insulin release (Akalestou et al., 2021), collectively accelerating insulin secretion and glucose clearance. The enhanced post-surgical glucose-insulin response, earlier peak timing, and faster glucose normalization offer significant benefits, as hyperglycemia is linked to vascular dysfunction, cardiovascular risk, oxidative stress, inflammation, and cognitive decline (Di Giuseppe et al., 2023). Interestingly, earlier timing of the glucose-insulin peaks in OGTT also seems to indicate the effectiveness of the antidiabetic drugs (Tran et al., 2018) and reflect improved insulin sensitivity and secretion in T2DM (Wang et al., 2018b), while delayed glucose peaks reflect pancreatic beta-cell dysfunction (Tran et al., 2018). In addition, increasing OGTT curve complexity is associated with better glucose tolerance (Tura et al., 2011). However, sharp glycemic peaks after OGTT have also been linked to atherosclerosis (Temelkova-Kurktschiev et al., 2000) and higher pulse pressure (Anan et al., 2008), suggesting potential risks in certain metabolic conditions. The metabolic impact of such peaks in MMT remains unexplored. The metabolic and predictive value of the sharp and early glucose-insulin peaks will be better understood in the future.

In this study, RYGB resulted in a higher peak glucose-insulin response, larger AUC, and prolonged glucose levels above 7.0 mmol/l compared to OAGB, though both procedures showed similar peak timing and return to baseline. This difference was particularly evident in MMT post-surgery. While earlier non-randomized studies suggested a superior glucose profile after OAGB due to differences in bypassed intestine length (Wang et al., 2018a), randomized controlled trials (RCTs) with comparable anatomy have found RYGB and OAGB to be metabolically equivalent at 1- and 2-year follow-ups (Heinonen et al., 2023; Robert et al., 2019). The RCT on the same data as in this study reports similar improvements in fasting glucose, Hba1c, and postprandial glucose and insulin AUCs in OGTT (Heinonen et al., 2023) as well as comparable glycemic responses to carbohydrate intake measured by continuous glucose monitoring (Ashrafi et al., 2021). Our findings suggest that the more pronounced glucose-insulin peak in RYGB stems from its anatomical differences, where nutrients reach key gut areas more rapidly, leading to faster glucose absorption and a sharper insulin response compared to OAGB. This accelerated glycemic response may enhance early glucose control. In contrast, OAGB's more gradual glucose absorption presents a smoother curve and may thus lower post-bariatric hypoglycemia risk, making it a preferable option for participants with preserved beta-cell function or milder metabolic disturbances. Although higher glycemic peaks in RYGB are linked to vascular risks like atherosclerosis and hypertension (Temelkova-Kurktschiev et al., 2000; Anan et al., 2008), studies indicate better hypertension improvement and reduced antihypertensive medication use post-RYGB (Heinonen et al., 2023). Given these dynamics, OAGB could, however, be a safer choice for individuals with a history of cardiovascular disease. Nonetheless, bariatric surgery remains a powerful metabolic intervention, providing substantial glycemic benefits regardless of the procedure (Sinclair et al., 2018).

Men exhibited higher peak glucose in OGTT compared to women, higher peak glucose and insulin, and longer TiH post-operatively in both OGTT and MMT, suggesting a more unhealthy profile in obesity and after weight loss. Typically, women have lower fasting plasma glucose, higher 2-hour glucose, and greater increases in 2-hour glucose following OGTT than men, though HbA1c levels are reported comparable between sexes (Sicree et al., 2008). Insulin sensitivity is also generally better in women (Nuutila et al., 1995). Research on obesity is sparse, with older studies indicating

that men have higher peak glucose at 1 hour during OGTT, but women surpass them at 2-3 hours (Young et al., 1979). Recent data shows that bariatric surgery leads to similar T2DM resolution between sexes, though only in men, AUC of insulin during OGTT predicts weight loss (Risi et al., 2022; Shu et al., 2024). Previous studies have not compared OGTT and MMT responses before and after bariatric surgery and between sexes. We confirm a blunted glucose-insulin response in both sexes in obesity, an unhealthier glucose profile for men throughout the weight loss, and sex-related advantages with sharper and quicker glucose-insulin responses post-operatively in favor of women.

Bariatric surgery results in excellent rates of T2DM remission (Courcoulas et al., 2024), but its effects on postprandial glycemia between persons with or without T2DM are not yet well characterized. Some studies suggest less weight loss in persons with T2DM (Rebelos et al., 2023) but do not assess postprandial glycemic response. In our study, individuals with T2DM lost more weight and showed better metabolic improvements but had higher glucose, lower insulin responses, and a flatter curve at all visits compared to non-diabetic individuals. This altered response reflects significant changes in glucose and insulin dynamics in T2DM, warranting more individualized characterization of postprandial glycemic responses.

The strengths of this study include the standardized OGTT and MMT measurements on the same individuals undergoing bariatric surgery, with follow-up extending to 6th and 12th months. The used cohort and its subgroups are representative of a general population undergoing bariatric surgery (Heinonen et al., 2023; Saarinen et al., 2019) making the results generalizable. The comparisons between RYGB and OAGB are from a randomized controlled trial designed to compare these procedures, using identical bypass limb lengths and equal stratification of participants by sex, age, and T2DM status. These factors enhance the reliability of comparisons. Additionally, the MMT's extended follow-up period of 360 minutes is a unique strength of this study. Analyzing responses across time and in subgroups further improves relevance by identifying procedure-specific, sex-related, and diabetes-associated patterns. While a larger, more diverse cohort would allow finer subgroup analyses, this study still provides valuable insights. Future research with an expanded dataset across different ethnicities could validate these findings. It is also worth noting that OGTT and MMT were performed on consecutive days, which may introduce slight day-to-day variation.

5 Conclusions

Utilizing Hierarchical Multi-Output Gaussian Process (HMOGP) regression modeling to analyze postprandial OGTT and MMT responses presents an advanced approach to studying postprandial glucose-insulin responses in obesity and weight loss. The employed HMOGP regression is capable of estimating individual-level effects — the personalized treatment response — in addition to the effects at the population level while accounting for and correcting for measurement errors and thus improving accuracy in complex multivariate data. This method not only unveils differences between metabolic cohorts after bariatric surgery but also a differential response between two operation types and sexes. This can lead to new knowledge in choosing the right treatments for participants and a more accurate modeling method for assessing OGTT and MMT responses in the clinic. Our study suggests that based on postprandial glucose dynamics, RYGB may be favorable over OAGB in hyperglycemic patients without risk for hypoglycemia or vascular diseases, where OAGB could be preferred. Thus, HMOGP regression offers a versatile tool for investigating and predicting glycemic responses in metabolic conditions, paving the way for personalized interventions in metabolic health.

Acknowledgments

We thank the participants for their invaluable time and effort. We also thank the study nurses, dieticians, and laboratory personnel for their help.

Funding

The study was supported by the Research Council of Finland (272376, 266286, 314383, and 335443 to KHP, 314457 to AJ, 361956 and 338417 to SH), the Finnish Medical Foundation (KHP, SH and AJ), the Finnish Diabetes Research Foundation (SH and KHP), the Orion Foundation (SH), the Novo Nordisk Foundation (NNF10OC1013354, NNF17OC0027232, and NNF20OC0060547 to KHP. NNF23SA0083953 for SH), European Foundation for the Diabetes Research (SH), the Paulo Foundation (SH), the Gyllenberg Foundation (KHP), Paavo Nurmi Foundation (SH), the Sigrid Juselius Foundation (KHP), Helsinki University Hospital Research Funds (SH, KHP, AJ), Government Research Funds (KHP, SH), the Research Council of Finland (Flagship programme: Finnish Center for Artificial Intelligence FCAI, and grants 352986, 358246 to PM) and EU (H2020 grant 101016775 and NextGenerationEU to PM).

Competing interests

The authors have no conflicts of interest to declare.

Data sharing statement

Individual participant data are not directly available but may be available after study close upon reasonable request from the authors. According to the General Data Protection Regulation of the European Union (679/2016), the principles of data protection should apply to any information concerning an identified or identifiable natural person and that personal data which have undergone pseudonymization, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. Thus, according to the GDPR, all pseudonymized data are considered personal data and cannot be published openly. Therefore, we are bound to the law and to the strict hospital policies, and are unable to share the data directly. However, the institutional (Helsinki and Uusimaa Hospital District) contact details for potential future data requests are as follows: `https://huspalvelu.microsoftcrmportals.com/fi-FI/`.

References

Adams TD, Davidson LE, Litwin SE, Kolotkin RL, LaMonte MJ, Pendleton RC, Strong MB, Vinik R, Wanner NA, Hopkins PN, et al. Health benefits of gastric bypass surgery after 6 years. Jama. 2012; 308(11):1122–1131.

Akalestou E, Suba K, Lopez-Noriega L, Georgiadou E, Chabosseau P, Gallie A, Wretlind A, Legido-Quigley C, Leclerc I, Salem V, et al. Intravital imaging of islet Ca2+ dynamics reveals enhanced β cell connectivity after bariatric surgery in mice. Nature Communications. 2021; 12(1):5165.

Almalki OM, Lee WJ, Chong K, Ser KH, Lee YC, Chen SC. Laparoscopic gastric bypass for the treatment of type 2 diabetes: a comparison of Roux-en-Y versus single anastomosis gastric bypass. Surgery for Obesity and Related Diseases. 2018; 14(4):509–515.

Álvarez M, Luengo D, Titsias M, Lawrence ND. Efficient multioutput Gaussian processes through variational inducing kernels. In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics* JMLR Workshop and Conference Proceedings; 2010. p. 25–32.

Alvarez MA, Lawrence ND. Computationally efficient convolved multiple output Gaussian processes. The Journal of Machine Learning Research. 2011; 12:1459–1500.

Anan F, Masaki T, Eto T, Fukunaga N, Iwao T, Kaneda K, Eshima N, Saikawa T, Yoshimatsu H. Postchallenge plasma glucose and glycemic spikes are associated with pulse pressure in patients with impaired glucose tolerance and essential hypertension. Hypertension Research. 2008; 31(8):1565–1571.

Ashrafi RA, Ahola AJ, Rosengård-Bärlund M, Saarinen T, Heinonen S, Juuti A, Marttinen P, Pietiläinen KH. Computational modelling of self-reported dietary carbohydrate intake on glucose concentrations in patients undergoing Roux-en-Y gastric bypass versus one-anastomosis gastric bypass. Annals of medicine. 2021; 53(1):1885–1895.

Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker MA, Guo J, Li P, Riddell A. Stan: A probabilistic programming language. Journal of statistical software. 2017; 76.

Cavin JB, Couvelard A, Lebtahi R, Ducroc R, Arapis K, Voitellier E, Cluzeaud F, Gillard L, Hourseau M, Mikail N, et al. Differences in alimentary glucose absorption and intestinal disposal of blood glucose after Roux-en-Y gastric bypass vs sleeve gastrectomy. Gastroenterology. 2016; 150(2):454–464.

Cersosimo E, Solis-Herrera C, E Trautmann M, Malloy J, L Triplitt C. Assessment of pancreatic β-cell function: review of methods and clinical applications. Current diabetes reviews. 2014; 10(1):2–42.

Cho W, Kim Y, Park J. Hierarchical anomaly detection using a multioutput Gaussian process. IEEE Transactions on Automation Science and Engineering. 2019; 17(1):261–272.

Courcoulas AP, Patti ME, Hu B, Arterburn DE, Simonson DC, Gourash WF, Jakicic JM, Vernon AH, Beck GJ, Schauer PR, et al. Long-term outcomes of medical management vs bariatric surgery in type 2 diabetes. JAMA. 2024; 331(8):654–664.

Cuesta Ramírez JJ, Álvarez López MA, editor, Hierarchical and Multiple Output Gaussian Process for the analysis of Colombian Cellular Networks. Pereira: Universidad Tecnológica de Pereira; 2015. `https://hdl.handle.net/11059/5968`.

Cummings DE, Overduin J, Foster-Schubert KE, Carlson MJ. Role of the bypassed proximal intestine in the anti-diabetic effects of bariatric surgery. Surgery for Obesity and Related Diseases. 2007; 3(2):109–115.

Dalla Man C, Rizza RA, Cobelli C. Meal simulation model of the glucose-insulin system. IEEE Transactions on biomedical engineering. 2007; 54(10):1740–1749.

Di Giuseppe G, Ciccarelli G, Soldovieri L, Capece U, Cefalo CM, Moffa S, Nista EC, Brunetti M, Cinti F, Gasbarrini A, et al. First-phase insulin secretion: can its evaluation direct therapeutic approaches? Trends in Endocrinology & Metabolism. 2023; 34(4):216–230.

Disse E, Pasquer A, Espalieu P, Poncet G, Gouillat C, Robert M. Greater weight loss with the omega loop bypass compared to the Roux-en-Y gastric bypass: a comparative study. Obesity surgery. 2014; 24:841–846.

Erdős B, van Sloun B, Adriaens ME, O'Donovan SD, Langin D, Astrup A, Blaak EE, Arts IC, van Riel NA. Personalized computational model quantifies heterogeneity in postprandial responses to oral glucose challenge. PLoS computational biology. 2021; 17(3):e1008852.

Ferrannini E, Mingrone G. Impact of different bariatric surgical procedures on insulin action and β-cell function in type 2 diabetes. Diabetes care. 2009; 32(3):514.

Garrish J, Chan C, Nychka D, Diniz Behn C. A Gaussian Process Model for Insulin Secretion Reconstruction with Uncertainty Quantification: Applications in Cystic Fibrosis. SIAM Journal on Applied Mathematics. 2023; p. S65–S81.

Gelman A, Carlin J, Stern H, Dunson D, Vehtari A, Rubin D. Bayesian Data Analysis. 3rd ed ed. United States: CreateSpace; 2013.

Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. BMC public health. 2009; 9:1–20.

Heinonen S, Buzkova J, Muniandy M, Kaksonen R, Ollikainen M, Ismail K, Hakkarainen A, Lundbom J, Lundbom N, Vuolteenaho K, et al. Impaired mitochondrial biogenesis in adipose tissue in acquired obesity. Diabetes. 2015; 64(9):3135–3145.

Heinonen S, Saarinen T, Meriläinen S, Sammalkorpi H, Penttilä AK, Koivikko M, Siira P, Karppinen J, Säiläkivi U, Rosengård-Bärlund M, et al. Roux-en-Y versus one-anastomosis gastric bypass (RYSA study): weight loss, metabolic improvements, and nutrition at 1 year after surgery, a multicenter randomized controlled trial. Obesity. 2023; 31(12):2909–2923.

Li J, Huang J, Zheng L, Li X. Application of artificial intelligence in diabetes education and management: present status and promising prospect. Frontiers in public health. 2020; 8:521222.

Li P, Chen S. Hierarchical gaussian processes model for multi-task learning. Pattern Recognition. 2018; 74:134–144.

Liu H, Cai J, Ong YS. Remarks on multi-output Gaussian process regression. Knowledge-Based Systems. 2018; 144:102–121.

MacKay DJ, et al. Introduction to Gaussian processes. NATO ASI series F computer and systems sciences. 1998; 168:133–166.

Magouliotis D, Tasiopoulou V, Tzovaras G. One anastomosis gastric bypass versus Roux-en-Y gastric bypass for morbid obesity: a meta-analysis. Clinical obesity. 2018; 8(3):159–169.

Mauvais-Jarvis F. Gender differences in glucose homeostasis and diabetes. Physiology & behavior. 2018; 187:20–23.

Moreno-Muñoz P, Artés A, Alvarez M. Heterogeneous multi-output Gaussian process prediction. Advances in neural information processing systems. 2018; 31.

Nuutila P, Knuuti MJ, Mäki M, Laine H, Ruotsalainen U, Teräs M, Haaparanta M, Solin O, Yki-Järvinen H. Gender and insulin sensitivity in the heart and in skeletal muscles: studies using positron emission tomography. Diabetes. 1995; 44(1):31–36.

Park S, Choi S. Hierarchical Gaussian process regression. In: *Proceedings of 2nd Asian conference on machine learning* JMLR Workshop and Conference Proceedings; 2010. p. 95–110.

Rebelos E, Moriconi D, Honka MJ, Anselmino M, Nannipieri M. Decreased weight loss following bariatric surgery in patients with type 2 diabetes. Obesity surgery. 2023; 33(1):179–187.

Risi R, Rossini G, Tozzi R, Pieralice S, Monte L, Masi D, Castagneto-Gissey L, Gallo IF, Strigari L, Casella G, et al. Sex difference in the safety and efficacy of bariatric procedures: a systematic review and meta-analysis. Surgery for Obesity and Related Diseases. 2022; 18(7):983–996.

Robert M, Espalieu P, Pelascini E, Caiazzo R, Sterkers A, Khamphommala L, Poghosyan T, Chevallier JM, Malherbe V, Chouillard E, et al. Efficacy and safety of one anastomosis gastric bypass versus Roux-en-Y gastric bypass for obesity (YOMEGA): a multicentre, randomised, open-label, non-inferiority trial. The Lancet. 2019; 393(10178):1299–1309.

Saarinen T, Meriläinen S, Koivukangas V, Pietiläinen KH, Juuti A. Prospective randomized controlled trial comparing the efficacy and safety of Roux-en-Y gastric bypass and one-anastomosis gastric bypass (the RYSA trial): trial protocol and interim analysis. Trials. 2019; 20:1–10.

Sandoval DA, Patti ME. Glucose metabolism after bariatric surgery: implications for T2DM remission and hypoglycaemia. Nature Reviews Endocrinology. 2023; 19(3):164–176.

Shu J, Zhu T, Xiong S, Liu T, Zhao Y, Huang X, Liu S. Sex dimorphism in the effect and predictors of weight loss after sleeve gastrectomy. Frontiers in Endocrinology. 2024; 14:1333051.

Sicree R, Zimmet P, Dunstan D, Cameron A, Welborn T, Shaw J. Differences in height explain gender differences in the response to the oral glucose tolerance test—the AusDiab study. Diabetic Medicine. 2008; 25(3):296–302.

Sinclair P, Docherty N, le Roux CW. Metabolic effects of bariatric surgery. Clinical chemistry. 2018; 64(1):72–81.

Sjöström L. Review of the key results from the Swedish Obese Subjects (SOS) trial–a prospective controlled intervention study of bariatric surgery. Journal of internal medicine. 2013; 273(3):219–234.

Temelkova-Kurktschiev TS, Koehler C, Henkel E, Leonhardt W, Fuecker K, Hanefeld M. Postchallenge plasma glucose and glycemic spikes are more strongly associated with atherosclerosis than fasting glucose or HbA1c level. Diabetes care. 2000; 23(12):1830–1834.

Tran S, Kramer CK, Zinman B, Choi H, Retnakaran R. Effect of chronic liraglutide therapy and its withdrawal on time to postchallenge peak glucose in type 2 diabetes. American Journal of Physiology-Endocrinology and Metabolism. 2018; 314(3):E287–E295.

Tura A, Morbiducci U, Sbrignadello S, Winhofer Y, Pacini G, Kautzky-Willer A. Shape of glucose, insulin, C-peptide curves during a 3-h oral glucose tolerance test: any relationship with the degree of glucose tolerance? American Journal of Physiology-Regulatory, Integrative and Comparative Physiology. 2011; 300(4):R941–R948.

Wang FG, Yan WM, Yan M, Song MM. Outcomes of Mini vs Roux-en-Y gastric bypass: a meta-analysis and systematic review. International Journal of Surgery. 2018; 56:7–14.

Wang X, Zhao X, Zhou R, Gu Y, Zhu X, Tang Z, Yuan X, Chen W, Zhang R, Qian C, et al. Delay in glucose peak time during the oral glucose tolerance test as an indicator of insulin resistance and insulin secretion in type 2 diabetes patients. Journal of diabetes investigation. 2018; 9(6):1288–1295.

Williams CK, Rasmussen CE. Gaussian processes for machine learning, vol. 2. MIT press Cambridge, MA; 2006.

Young RL, Fuchs RJ, Woltjen MJ, Cooperman MT. Glucose-insulin response to oral glucose in a healthy obese population. Diabetes. 1979; 28(3):208–212.

Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, Ben-Yacov O, Lador D, Avnit-Sagi T, Lotan-Pompan M, et al. Personalized nutrition by prediction of glycemic responses. Cell. 2015; 163(5):1079–1094.

Zimmet P, Magliano D, Matsuzawa Y, Alberti G, Shaw J. The metabolic syndrome: a global public health problem and a new definition. Journal of atherosclerosis and thrombosis. 2005; 12(6):295–300.

A Modeling Details

We normalized the data in both time and value domains before the training, separately for insulin and glucose values. We used a fixed lengthscale parameters for all GPs, which is set to be $120/\sigma_{\text{time}}$, where σ_{time} is the standard deviation of the sample times. We used the Matern52 kernel for all GPs. The main difference between individual-level GPs and the population-level GPs was the prior value of the kernel variance. We used $InverseGamma(1, 1)$ distribution for the population level GPs, while for the individual level GPs, we used a stronger $InverseGamma(1, 0.1)$ distribution. This setup allowed the model to learn the shared behaviors with the population-level GPs while the individual-level GPs learned individual differences. For measurement error σ, we used $InverseGamma(1, 0.01)$ as a prior distribution. For $\boldsymbol{\mu}^{\text{population}}$, we used $\mathcal{N}(0, 1)$ as a prior distribution, and for $\boldsymbol{\mu}_n^{\text{patient}}$ we used $\mathcal{N}(\boldsymbol{\mu}^{\text{population}}, \sigma_n^{\text{patient}})$ as a prior distribution, where $\sigma_n^{\text{patient}} \sim InverseGamma(1, 0.5)$. The rest of the details and full Stan code are available at `https://github.com/onurpoyraz/HMOGP`.

B Individual Responses

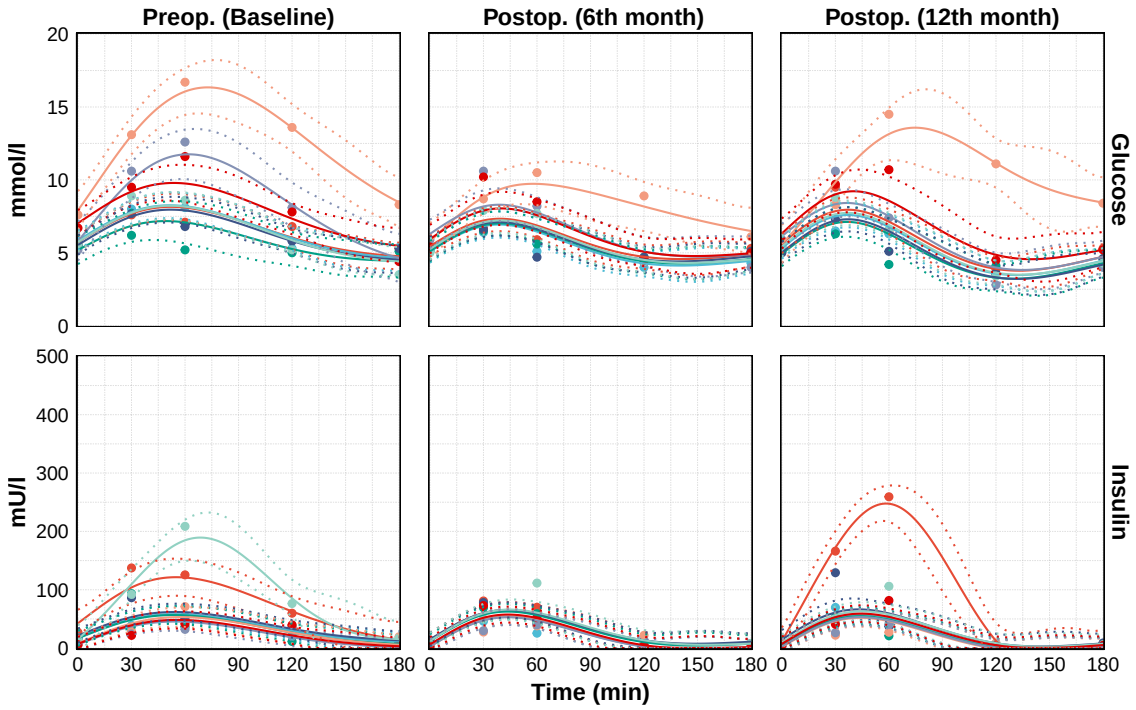

Figure B.1 | Response of the glucose and insulin levels and predicted model fit for different individuals with applied OGTT test.

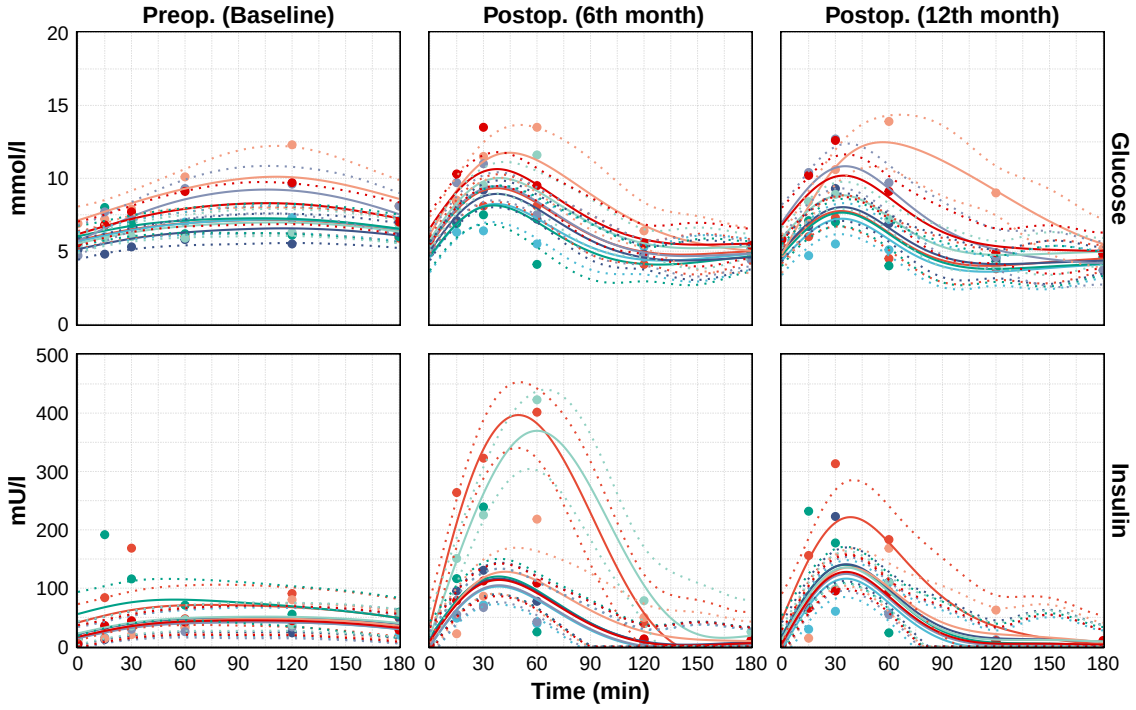

Figure B.2 | Response of the glucose and insulin levels and predicted model fit for different individuals with applied MMT test.

C MMT Model Fitting

MMT includes additional data which sampled at 15th, 240th, and 360th minutes compared to OGTT. So, for an healthy comparison we trained the model on MMT using the same datapoints exist on the OGTT (partial data). Figure C.1 compares the differences between the model fits with the full data and partial data on MMT.

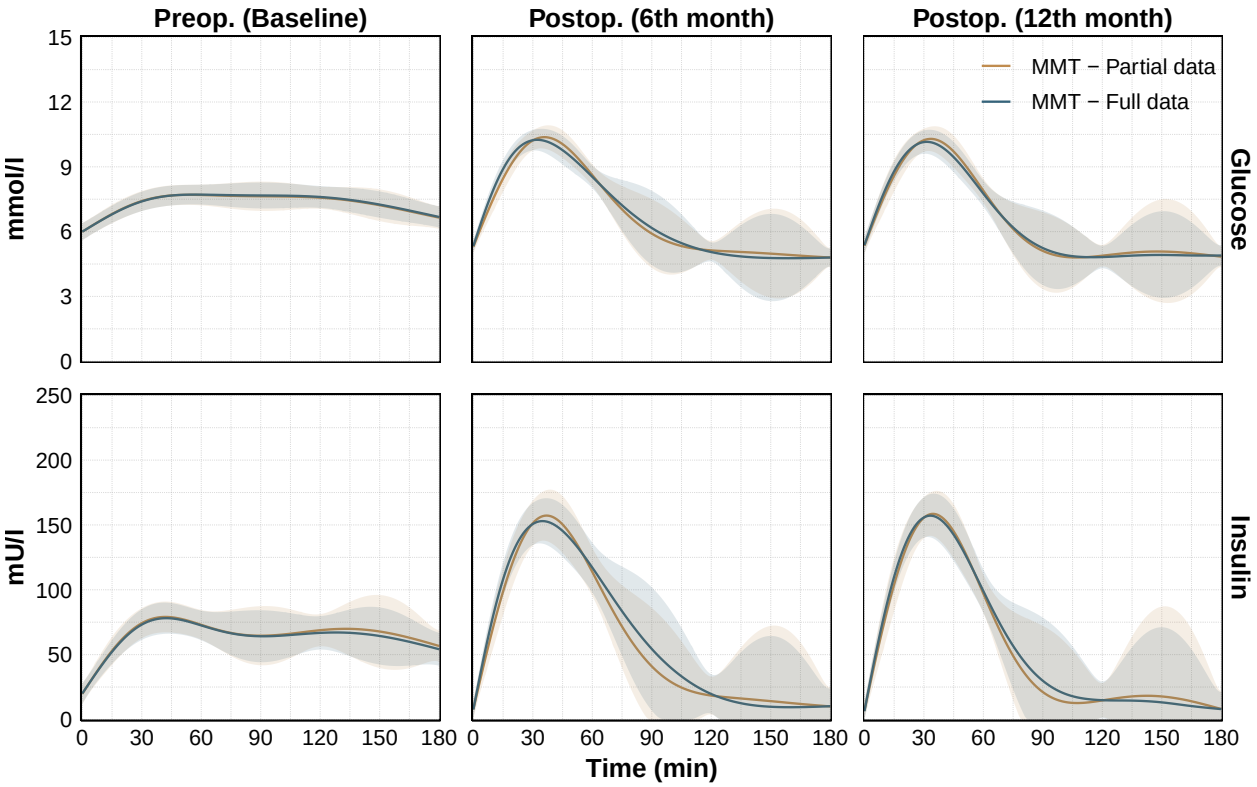

Figure C.1 | Difference of model predictions with partial data vs. full data on the MMT. Shaded areas represent 95% confidence intervals (CI).

D Comparison by Diabetes

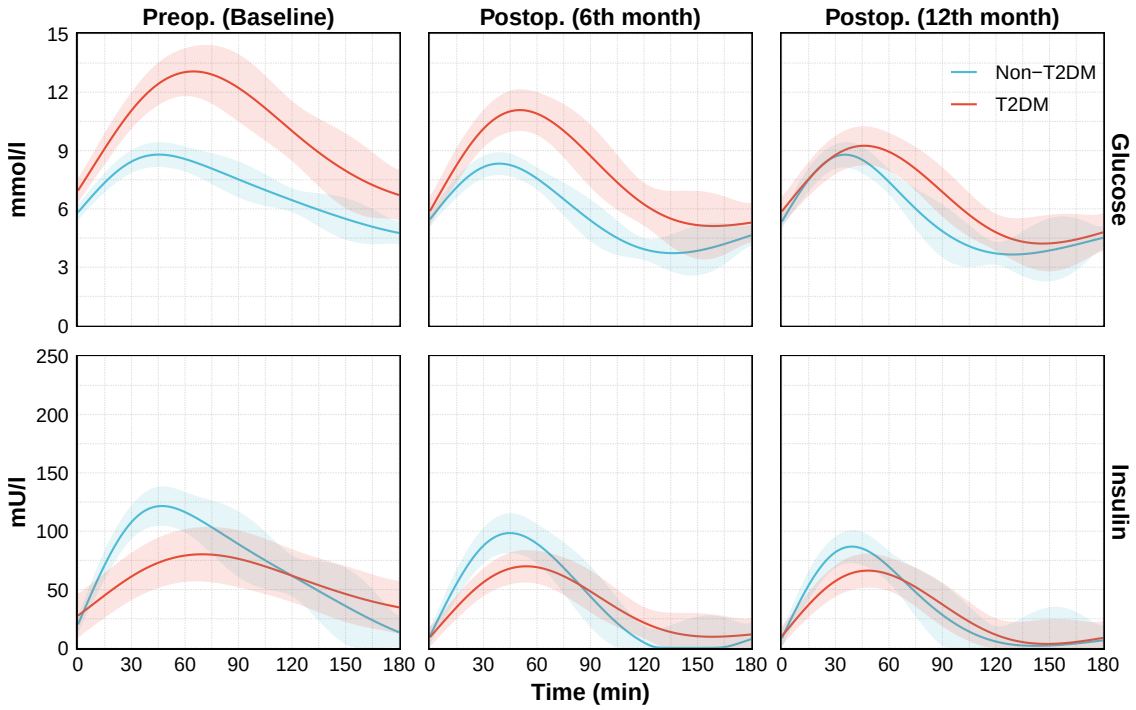

Figure D.1 | The comparison of model predicted average glucose and insulin response of the Non-T2DM and T2DM participants in OGTT test. Shaded areas represent 95% confidence intervals (CI).

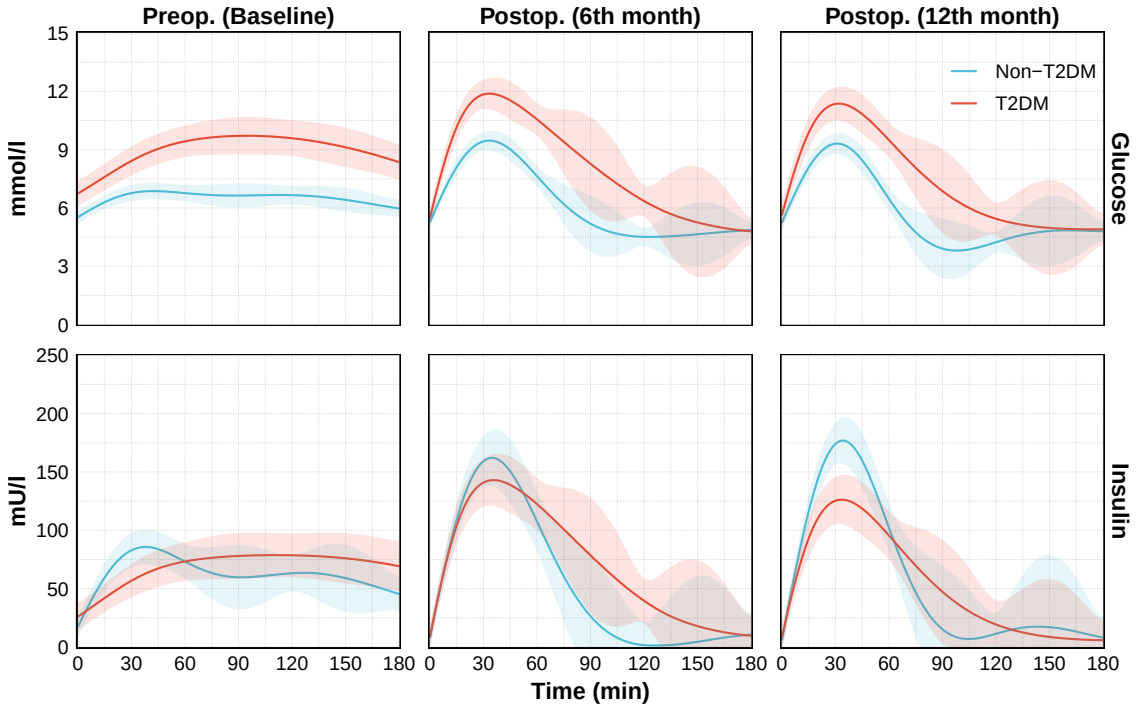

Figure D.2 | The comparison of model predicted average glucose and insulin response of the Non-T2DM and T2DM participants in MMT test. Shaded areas represent 95% confidence intervals (CI).

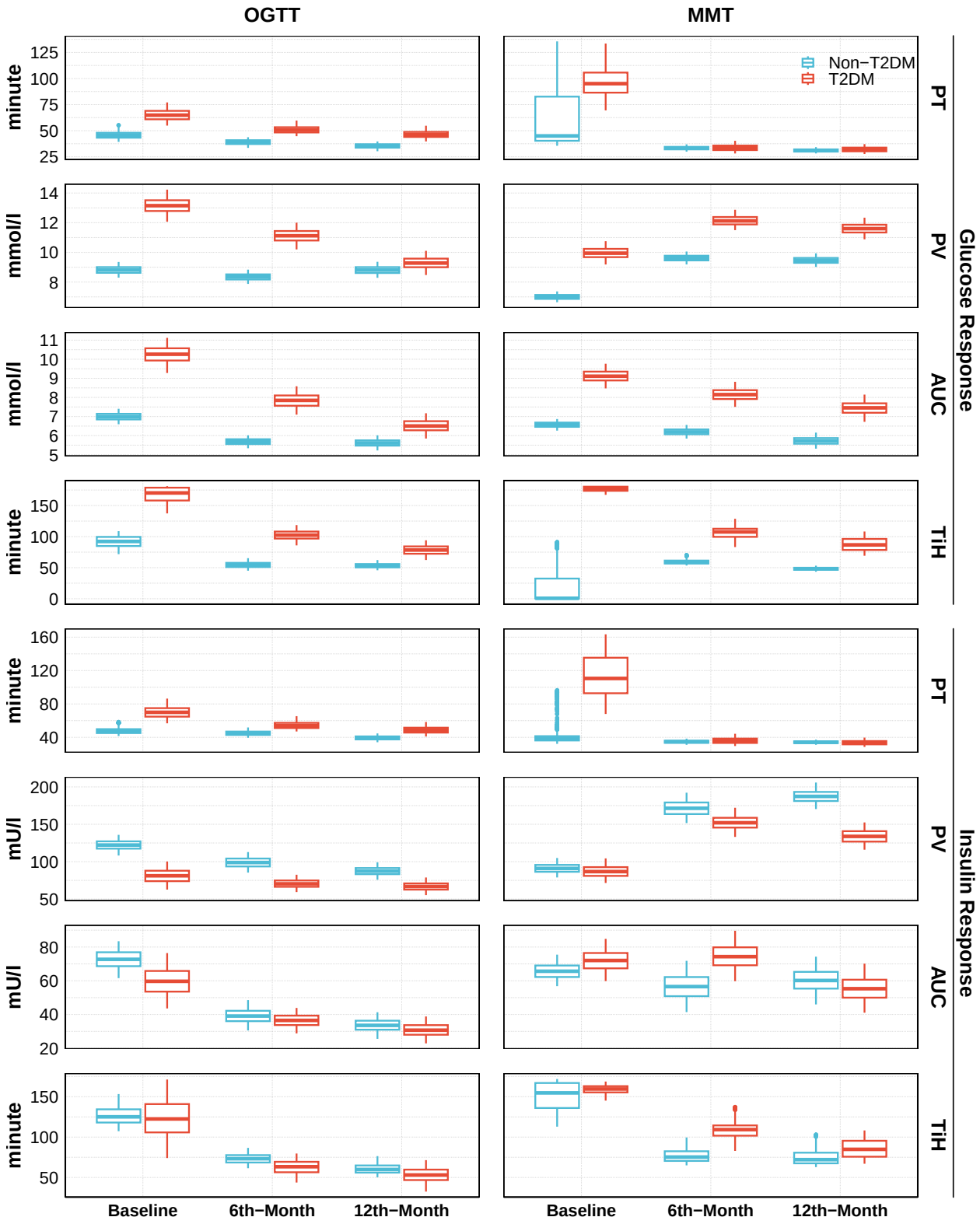

Figure D.3 | The box-plot of each metric for T2DM comparison. See Appendix F for the 95% CIs and p-values.

E Comparison by Sex

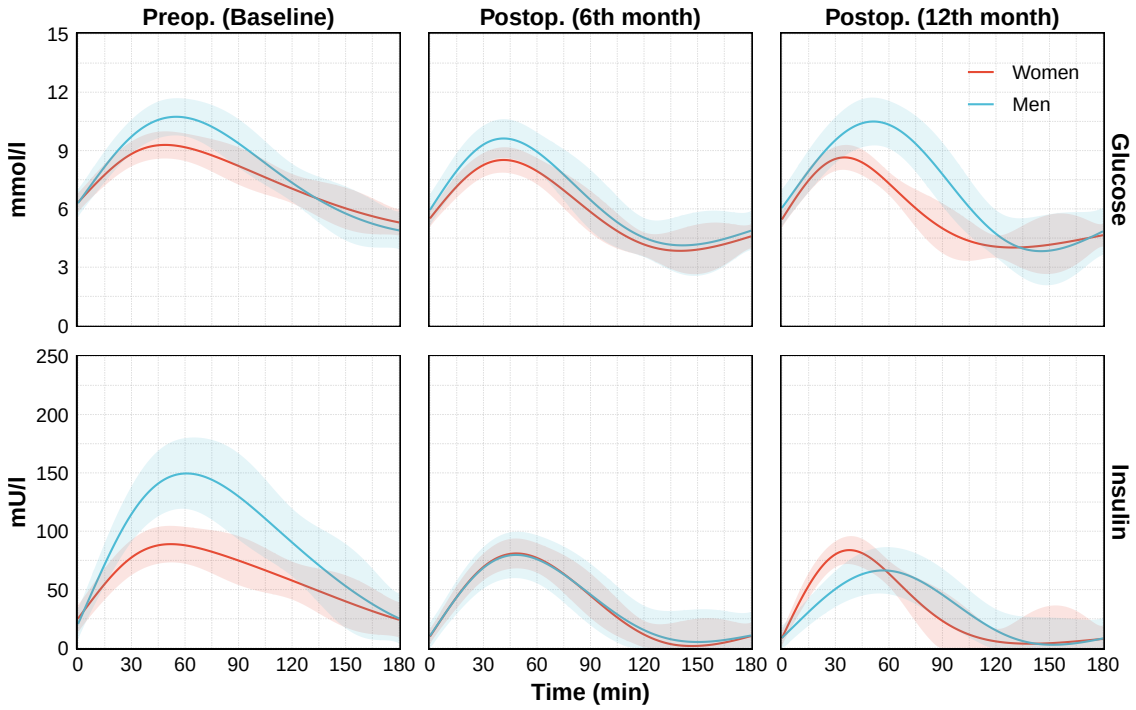

Figure E.1 | The comparison of model predicted average glucose and insulin response of the women and men participants in OGTT test. Shaded areas represent 95% confidence intervals (CI).

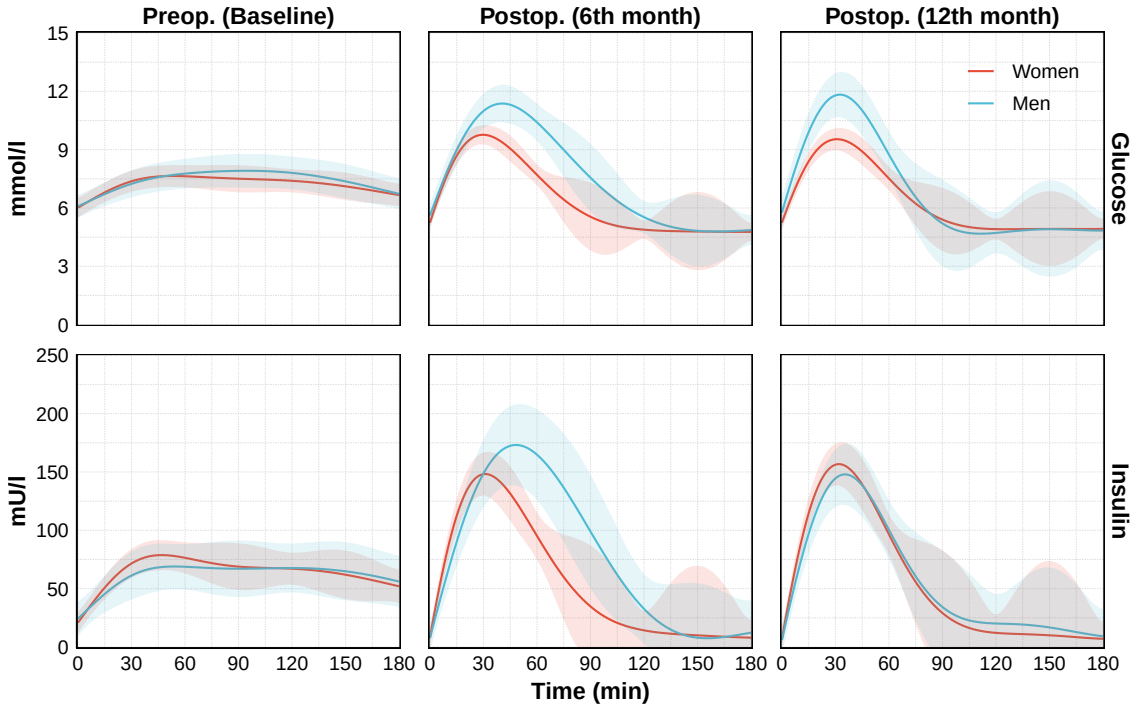

Figure E.2 | The comparison of model predicted average glucose and insulin response of the women and men participants in MMT test.

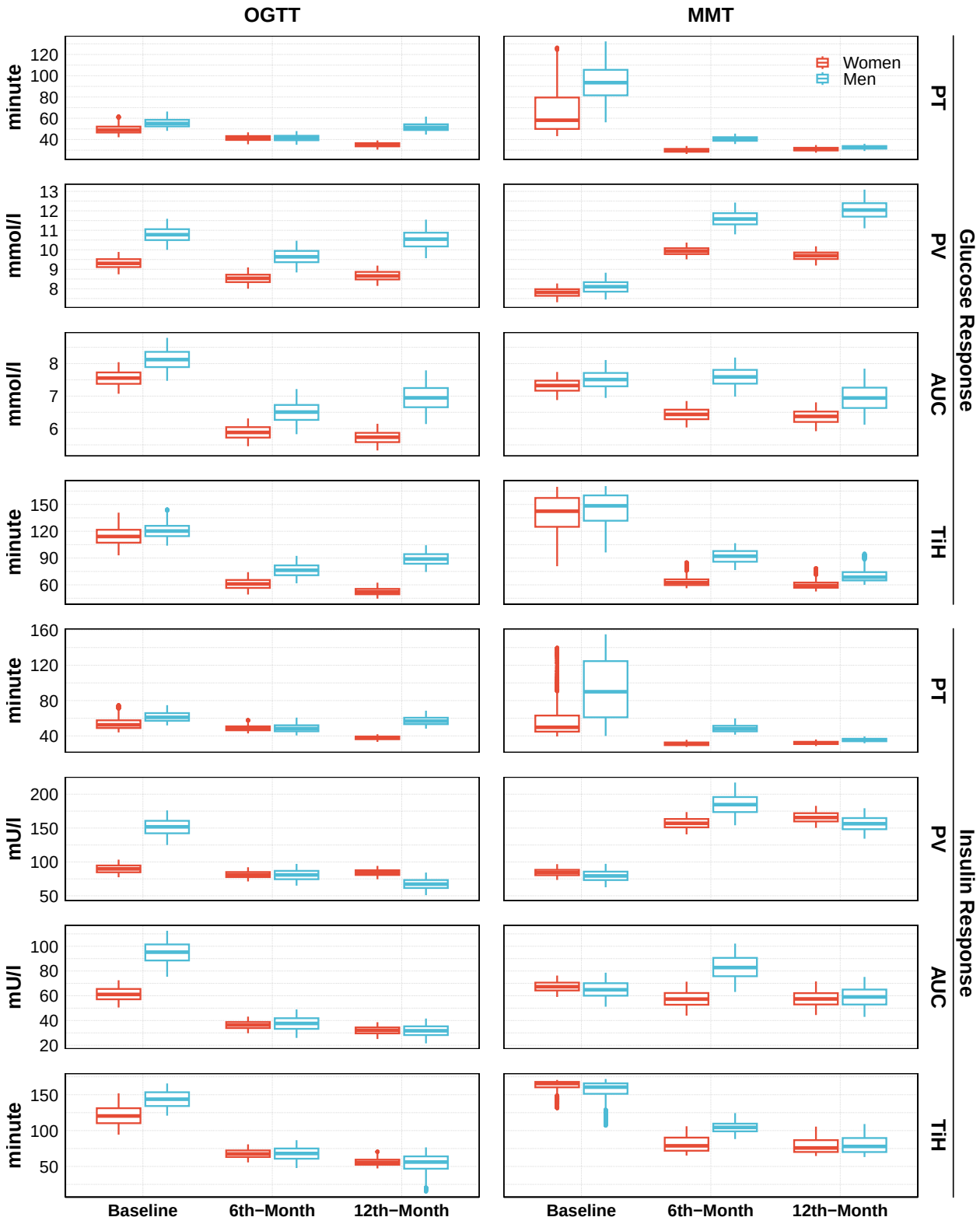

Figure E.3 | The box-plot of each metric for sex comparison. See Appendix F for the 95% CIs and p-values.

F Evaluation of Metrics

Table F.1 | Peaking time (PT, minute) of the glucose and insulin responses with 95% confidence intervals and p-values. p-values (except $p(\text{B}\rightarrow12)$) shows the statistical significance of the difference between adjacent cells. $p(\text{B}\rightarrow12)$ represents the statistical significance of the difference between Baseline and 12th month values.

Peaking Time (PT)				**Baseline**	$p(\text{B}\rightarrow6)$	**6th month**	$p(\text{B}\rightarrow12)$	**12th month**
Glucose Response	**Test**	**Overall**	**OGTT**	50 (45-62)	0.013	41 (35-45)	0.001	38 (33-42)
			p	0.199	-	0.027	-	0.038
			MMT	77 (44-129)	0.001	32 (28-37)	<0.001	31 (29-35)
	Operation	**OGTT**	**RYGB**	52 (45-69)	0.018	41 (35-46)	0.002	39 (34-43)
			p	0.419	-	0.335	-	0.488
			OAGB	51 (43-61)	0.110	43 (37-50)	0.029	39 (32-45)
		MMT	**RYGB**	84 (47-133)	0.001	33 (28-38)	<0.001	31 (28-35)
			p	0.388	-	0.492	-	0.456
			OAGB	70 (45-125)	<0.001	33 (28-37)	<0.001	32 (28-35)
	Diabetes	**OGTT**	**Non-T2DM**	46 (39-55)	0.083	39 (34-44)	0.019	35 (30-40)
			p	0.021	-	0.008	-	0.016
			T2DM	65 (55-77)	0.048	51 (45-60)	0.009	46 (40-55)
		MMT	**Non-T2DM**	45 (36-134)	0.026	33 (30-37)	0.006	31 (28-34)
			p	0.187	-	0.473	-	0.386
			T2DM	95 (69-136)	0.002	33 (28-40)	<0.001	32 (28-37)
	Sex	**OGTT**	**Women**	49 (42-62)	0.078	41 (36-47)	0.002	35 (30-39)
			p	0.208	-	0.495	-	0.001
			Men	55 (48-66)	0.012	41 (35-48)	0.287	51 (45-62)
		MMT	**Women**	58 (43-126)	0.001	30 (26-34)	<0.001	31 (28-35)
			p	0.206	-	0.002	-	0.306
			Men	93 (56-132)	0.002	41 (36-46)	<0.001	33 (29-36)
Insulin Response	**Test**	**Overall**	**OGTT**	53 (46-71)	0.262	49 (43-59)	0.018	42 (37-46)
			p	0.297	-	0.002	-	0.020
			MMT	44 (37-145)	0.041	35 (31-40)	0.017	33 (30-37)
	Operation	**OGTT**	**RYGB**	50 (42-69)	0.348	48 (42-56)	0.054	41 (36-46)
			p	0.149	-	0.374	-	0.217
			OAGB	62 (50-78)	0.120	50 (42-62)	0.033	45 (38-54)
		MMT	**RYGB**	44 (35-139)	0.274	40 (33-46)	0.021	32 (29-36)
			p	0.153	-	0.122	-	0.334
			OAGB	83 (46-147)	0.001	34 (30-38)	0.001	34 (30-38)
	Diabetes	**OGTT**	**Non-T2DM**	47 (42-58)	0.326	45 (39-52)	0.048	39 (34-45)
			p	0.027	-	0.067	-	0.052
			T2DM	70 (57-86)	0.068	54 (47-65)	0.015	48 (41-58)
		MMT	**Non-T2DM**	39 (32-97)	0.190	35 (31-38)	0.140	34 (31-37)
			p	0.045	-	0.407	-	0.444
			T2DM	111 (68-163)	0.003	36 (30-44)	0.001	34 (29-40)
	Sex	**OGTT**	**Women**	53 (44-75)	0.292	48 (43-58)	0.004	38 (33-42)
			p	0.217	-	0.478	-	0.003
			Men	61 (52-75)	0.076	48 (41-61)	0.306	57 (48-69)
		MMT	**Women**	50 (39-140)	0.002	31 (28-36)	0.001	32 (29-36)
			p	0.257	-	0.001	-	0.159
			Men	90 (40-155)	0.157	48 (41-60)	0.015	35 (32-39)

Table F.2 | Peak value (PV, mmol/l for glucose, mU/l for insulin) of the glucose and insulin responses with 95% confidence intervals and p-values. p-values (except $p(\mathrm{B}\rightarrow 12)$) shows the statistical significance of the difference between adjacent cells. $p(\mathrm{B}\rightarrow 12)$ represents the statistical significance of the difference between Baseline and 12th month values.

Peak Value (PV)				**Baseline**	$p(\mathrm{B}\rightarrow 6)$	**6th month**	$p(\mathrm{B}\rightarrow 12)$	**12th month**
Glucose Response	**Test**	**Overall**	**OGTT**	9.9 (9.4-10.4)	0.011	9.0 (8.5-9.4)	0.018	9.0 (8.5-9.5)
			p	<0.001	-	<0.001	-	0.004
			MMT	7.8 (7.4-8.2)	<0.001	10.3 (9.9-10.7)	<0.001	10.2 (9.7-10.6)
	Operation	**OGTT**	**RYGB**	10.0 (9.5-10.6)	0.020	9.0 (8.4-9.6)	0.474	10.0 (9.4-10.6)
			p	0.230	-	0.336	-	<0.001
			OAGB	9.6 (8.8-10.4)	0.102	8.8 (8.1-9.4)	0.004	7.8 (7.1-8.5)
		MMT	**RYGB**	7.8 (7.3-8.2)	<0.001	11.4 (10.9-11.9)	<0.001	11.7 (11.2-12.1)
			p	0.388	-	<0.001	-	<0.001
			OAGB	7.6 (7.0-8.2)	0.006	8.8 (8.3-9.3)	0.002	9.1 (8.5-9.7)
	Diabetes	**OGTT**	**Non-T2DM**	8.8 (8.3-9.4)	0.138	8.3 (7.9-8.8)	0.488	8.8 (8.3-9.4)
			p	<0.001	-	<0.001	-	0.214
			T2DM	13.1 (12.1-14.2)	0.010	11.1 (10.2-12.0)	<0.001	9.3 (8.5-10.1)
		MMT	**Non-T2DM**	6.9 (6.6-7.3)	<0.001	9.5 (9.1-9.9)	<0.001	9.3 (8.9-9.8)
			p	<0.001	-	<0.001	-	<0.001
			T2DM	9.8 (9.1-10.6)	<0.001	11.9 (11.3-12.6)	0.008	11.4 (10.7-12.1)
	Sex	**OGTT**	**Women**	9.3 (8.7-9.9)	0.051	8.5 (8.0-9.1)	0.087	8.7 (8.1-9.2)
			p	0.006	-	0.026	-	0.003
			Men	10.8 (10.0-11.6)	0.045	9.6 (8.8-10.5)	0.377	10.5 (9.6-11.5)
		MMT	**Women**	7.7 (7.2-8.2)	<0.001	9.8 (9.4-10.2)	<0.001	9.6 (9.1-10.0)
			p	0.280	-	0.002	-	<0.001
			Men	8.0 (7.4-8.7)	<0.001	11.4 (10.6-12.2)	<0.001	11.8 (10.9-12.8)
Insulin Response	**Test**	**Overall**	**OGTT**	97 (86-109)	0.054	83 (74-93)	0.016	79 (70-89)
			p	0.036	-	<0.001	-	<0.001
			MMT	80 (70-90)	<0.001	154 (139-168)	<0.001	158 (144-172)
	Operation	**OGTT**	**RYGB**	108 (94-123)	0.398	104 (91-119)	0.029	87 (75-100)
			p	0.062	-	<0.001	-	0.062
			OAGB	86 (68-106)	0.091	69 (55-81)	0.124	71 (61-83)
		MMT	**RYGB**	88 (76-101)	<0.001	217 (191-241)	<0.001	187 (169-204)
			p	0.040	-	<0.001	-	<0.001
			OAGB	70 (57-82)	<0.001	122 (106-138)	<0.001	126 (109-142)
	Diabetes	**OGTT**	**Non-T2DM**	122 (108-136)	0.028	99 (85-113)	<0.001	87 (76-99)
			p	0.006	-	0.004	-	0.020
			T2DM	81 (63-100)	0.218	70 (59-83)	0.140	67 (55-79)
		MMT	**Non-T2DM**	87 (75-100)	<0.001	162 (144-182)	<0.001	177 (161-195)
			p	0.360	-	0.120	-	<0.001
			T2DM	83 (68-99)	<0.001	144 (126-163)	0.001	127 (110-145)
	Sex	**OGTT**	**Women**	90 (77-103)	0.203	81 (71-92)	0.288	84 (74-94)
			p	<0.001	-	0.469	-	0.068
			Men	152 (125-176)	<0.001	81 (65-97)	<0.001	67 (51-84)
		MMT	**Women**	81 (70-92)	<0.001	149 (133-164)	<0.001	157 (142-173)
			p	0.340	-	0.087	-	0.281
			Men	76 (60-93)	<0.001	175 (146-205)	<0.001	148 (127-170)

Table F.3 | Time-normalized **Area Under Curve (AUC), mmol/l for glucose, mU/l for insulin)** of the glucose and insulin responses with 95% confidence intervals and p-values. p-values (except $p(\text{B}\rightarrow12)$) shows the statistical significance of the difference between adjacent cells. $p(\text{B}\rightarrow12)$ represents the statistical significance of the difference between Baseline and 12th month values.

Area Under Curve (AUC)				**Baseline**	$p(\text{B}\rightarrow6)$	**6th month**	$p(\text{B}\rightarrow12)$	**12th month**
Glucose Response	**Test**	**Overall**	**OGTT**	7.8 (7.4-8.2)	<0.001	6.1 (5.7-6.4)	<0.001	6.0 (5.6-6.4)
			p	0.098	-	0.026	-	0.117
			MMT	7.4(7.0-7.7)	0.033	6.8 (6.4-7.1)	0.004	6.4 (6.0-6.9)
	Operation	**OGTT**	**RYGB**	8.0 (7.5-8.4)	<0.001	6.1 (5.6-6.5)	<0.001	6.3 (5.8-6.8)
			p	0.129	-	0.460	-	0.094
			OAGB	7.5 (6.8-8.1)	0.004	6.1 (5.6-6.7)	<0.001	5.7 (5.1-6.3)
		MMT	**RYGB**	7.3 (6.8-7.6)	0.381	7.1 (6.7-7.6)	0.120	6.8 (6.3-7.3)
			p	0.494	-	0.012	-	0.156
			OAGB	7.2 (6.7-7.8)	0.013	6.3 (5.8-6.7)	0.034	6.3 (5.8-6.9)
	Diabetes	**OGTT**	**Non-T2DM**	7.0 (6.6-7.4)	<0.001	5.7 (5.3-6.0)	<0.001	5.6 (5.2-6.0)
			p	<0.001	-	<0.001	-	0.025
			T2DM	10.3 (9.3-11.1)	0.002	7.8 (7.1-8.6)	<0.001	6.5 (5.8-7.2)
		MMT	**Non-T2DM**	6.5 (6.2-6.8)	0.090	6.2 (5.8-6.5)	0.005	5.7 (5.3-6.1)
			p	<0.001	-	<0.001	-	<0.001
			T2DM	9.0 (8.4-9.6)	0.049	8.1 (7.4-8.7)	0.004	7.4 (6.7-8.1)
	Sex	**OGTT**	**Women**	7.6 (7.1-8.0)	<0.001	5.9 (5.5-6.3)	<0.001	5.7 (5.3-6.1)
			p	0.116	-	0.092	-	0.016
			Men	8.1 (7.5-8.8)	0.002	6.5 (5.8-7.2)	0.034	6.9 (6.1-7.8)
		MMT	**Women**	7.3 (6.8-7.7)	0.008	6.4 (6.0-6.8)	0.009	6.3 (5.9-6.8)
			p	0.326	-	0.008	-	0.167
			Men	7.4 (6.9-8.0)	0.428	7.5 (6.9-8.1)	0.189	6.9 (6.1-7.8)
Insulin Response	**Test**	**Overall**	**OGTT**	64 (55-73)	<0.001	38 (31-44)	<0.001	33 (27-39)
			p	0.490	-	<0.001	-	<0.001
			MMT	64 (57-72)	0.470	63 (51-75)	0.180	56 (45-68)
	Operation	**OGTT**	**RYGB**	68 (58-79)	0.008	46 (36-55)	<0.001	36 (28-44)
			p	0.200	-	0.040	-	0.203
			OAGB	58 (44-74)	0.006	32 (23-40)	0.002	30 (22-38)
		MMT	**RYGB**	68 (59-79)	0.020	96 (77-112)	0.336	64 (51-79)
			p	0.093	-	<0.001	-	0.048
			OAGB	57 (48-66)	0.162	48 (36-60)	0.104	46 (34-58)
	Diabetes	**OGTT**	**Non-T2DM**	73 (62-83)	<0.001	39 (31-49)	<0.001	34 (26-41)
			p	0.132	-	0.354	-	0.337
			T2DM	60 (44-76)	0.018	37 (29-44)	0.005	31 (23-39)
		MMT	**Non-T2DM**	63 (55-72)	0.197	54 (40-69)	0.289	58 (44-71)
			p	0.261	-	0.079	-	0.353
			T2DM	69 (57-81)	0.408	71 (57-85)	0.078	53 (40-67)
	Sex	**OGTT**	**Women**	61 (51-73)	<0.001	36 (30-43)	<0.001	32 (25-39)
			p	0.009	-	0.446	-	0.494
			Men	95 (75-112)	<0.001	38 (26-49)	<0.001	32 (22-42)
		MMT	**Women**	64 (57-73)	0.141	55 (42-68)	0.151	55 (43-68)
			p	0.400	-	0.042	-	0.454
			Men	62 (49-75)	0.107	79 (60-97)	0.316	57 (41-72)

Table F.4 | Time in the Risk Zone (TiH, minute), where hyperglycemia and hyperinsulinemia are defined as glucose > 7.0mmol/l and insulin > 50mU/l, respectively, with 95 % confidence intervals and p-values. p-values (except $p(\mathrm{B}\rightarrow 12)$) shows the statistical significance of the difference between adjacent cells. $p(\mathrm{B}\rightarrow 12)$ represents the statistical significance of the difference between Baseline and 12th month values.

Time in the Risk Zone (TiH)				**Baseline**	$p(\mathrm{B}\rightarrow 6)$	**6th month**	$p(\mathrm{B}\rightarrow 12)$	**12th month**
Glucose Response	**Test**	**Overall**	**OGTT**	116 (103-135)	<0.001	66 (56-78)	<0.001	63 (54-75)
			p	0.118	-	0.259	-	0.400
			MMT	144 (109-166)	0.006	72 (63-93)	0.004	61 (55-78)
	Operation	**OGTT**	**RYGB**	125 (110-147)	<0.001	67 (55-80)	<0.001	70 (61-85)
			p	0.088	-	0.500	-	0.040
			OAGB	105 (84-126)	0.007	67 (53-82)	<0.001	47 (20-66)
		MMT	**RYGB**	133 (84-162)	0.054	83 (70-101)	0.033	69 (62-91)
			p	0.458	-	0.014	-	0.056
			OAGB	136 (27-174)	0.090	53 (45-65)	0.091	54 (45-68)
	Diabetes	**OGTT**	**Non-T2DM**	92 (72-109)	0.004	54 (45-65)	<0.001	53 (46-62)
			p	<0.001	-	<0.001	-	0.016
			T2DM	170 (137-181)	0.004	102 (86-119)	<0.001	78 (62-94)
		MMT	**Non-T2DM**	2 (1-74)	0.082	58 (53-68)	0.116	47 (43-52)
			p	<0.001	-	0.001	-	0.001
			T2DM	176 (166-181)	<0.001	106 (82-126)	<0.001	85 (69-107)
	Sex	**OGTT**	**Women**	114 (93-141)	<0.001	61 (49-74)	<0.001	52 (45-62)
			p	0.359	-	0.096	-	0.001
			Men	120 (104-144)	<0.001	76 (62-92)	0.014	89 (74-104)
		MMT	**Women**	136 (65-168)	0.047	61 (55-83)	0.044	58 (52-75)
			p	0.442	-	0.026	-	0.141
			Men	144 (88-169)	0.059	91 (76-106)	0.028	68 (59-93)
Insulin Response	**Test**	**Overall**	**OGTT**	121 (102-148)	<0.001	70 (58-82)	<0.001	59 (50-74)
			p	0.053	-	0.081	-	0.071
			MMT	165 (127-170)	0.004	89 (72-107)	0.001	74 (64-102)
	Operation	**OGTT**	**RYGB**	125 (107-152)	0.005	80 (68-93)	<0.001	63 (53-79)
			p	0.301	-	0.036	-	0.253
			OAGB	112 (75-154)	0.022	57 (31-74)	0.014	55 (40-71)
		MMT	**RYGB**	168 (124-173)	0.018	108 (90-126)	0.004	80 (68-107)
			p	0.160	-	0.028	-	0.154
			OAGB	148 (83-163)	0.037	67 (57-97)	0.034	64 (55-96)
	Diabetes	**OGTT**	**Non-T2DM**	125 (107-153)	0.002	73 (61-87)	<0.001	60 (50-76)
			p	0.450	-	0.202	-	0.278
			T2DM	122 (74-171)	0.034	63 (44-80)	0.023	53 (33-71)
		MMT	**Non-T2DM**	148 (106-171)	0.009	74 (64-98)	0.009	71 (62-101)
			p	0.378	-	0.051	-	0.283
			T2DM	158 (141-167)	0.017	108 (82-132)	0.003	83 (66-106)
	Sex	**OGTT**	**Women**	120 (94-152)	0.004	67 (56-81)	0.002	56 (47-71)
			p	0.152	-	0.490	-	0.493
			Men	144 (121-166)	<0.001	68 (48-87)	<0.001	56 (15-77)
		MMT	**Women**	163 (125-169)	0.001	77 (64-105)	0.002	74 (64-104)
			p	0.403	-	0.102	-	0.484
			Men	158 (96-170)	0.075	103 (87-122)	0.033	76 (62-107)

G Data Statistics

Table G.1 | Metabolic features of the participants by different operations, sexes, and diabetes over each visit. **B** represents baseline visit, **6th** corresponds to 6th month visit, and **12th** shows 12th month visit.

Metabolic Features		Overall	OAGB	RYGB	Male	Female	T2DM	Non-T2DM
Age (years)	**B**	46.0 ± 7.0	46.6 ± 7.4	45.4 ± 6.6	50.7 ± 7.4	43.9 ± 5.8	44.3 ± 6.3	47.3 ± 7.3
	6th	46.5 ± 7.0	47.1 ± 7.4	45.9 ± 6.6	51.2 ± 7.4	44.4 ± 5.8	44.8 ± 6.3	47.8 ± 7.3
	12th	47.0 ± 7.0	47.6 ± 7.4	46.4 ± 6.6	51.7 ± 7.4	44.9 ± 5.8	45.3 ± 6.3	48.3 ± 7.3
Weight (kg)	**B**	130 ± 19	127 ± 15	135 ± 23	138 ± 22	127 ± 17	132 ± 21	130 ± 18
	6th	102 ± 16	99 ± 14	107 ± 19	109 ± 18	101 ± 16	101 ± 16	104 ± 17
	12th	98 ± 16	94 ± 14	103 ± 18	106 ± 19	95 ± 15	95 ± 16	101 ± 17
Bmi (kg/m2)	**B**	45 ± 6	44 ± 6	46 ± 6	42 ± 5	46 ± 6	46 ± 6	44 ± 6
	6th	35 ± 5	35 ± 6	36 ± 5	34 ± 5	37 ± 5	35 ± 5	36 ± 5
	12th	34 ± 5	33 ± 6	35 ± 5	33 ± 5	35 ± 5	34 ± 6	35 ± 5
Hba1c (mmol/mol)	**B**	39 ± 10	38 ± 6	41 ± 13	38 ± 7	40 ± 11	47 ± 12	34 ± 4
	6th	34 ± 5	34 ± 3	33 ± 6	34 ± 4	34 ± 5	37 ± 4	32 ± 4
	12th	34 ± 7	34 ± 3	35 ± 10	35 ± 6	34 ± 7	38 ± 8	31 ± 4
Cholesterol (mmol/l)	**B**	4.5 ± 0.9	4.5 ± 1.0	4.6 ± 0.7	4.3 ± 1.3	4.6 ± 0.7	4.5 ± 1.1	4.5 ± 0.8
	6th	3.8 ± 0.7	3.7 ± 0.7	3.9 ± 0.6	3.5 ± 0.8	3.9 ± 0.6	3.7 ± 0.8	3.9 ± 0.6
	12th	3.9 ± 1.0	3.8 ± 0.7	4.3 ± 1.2	3.8 ± 1.1	4.1 ± 0.9	3.8 ± 0.9	4.1 ± 1.1
HDL Cholesterol (mmol/l)	**B**	1.2 ± 0.2	1.2 ± 0.2	1.2 ± 0.2	1.2 ± 0.2	1.3 ± 0.2	1.2 ± 0.3	1.2 ± 0.2
	6th	1.2 ± 0.2	1.2 ± 0.2	1.3 ± 0.2	1.2 ± 0.2	1.2 ± 0.2	1.1 ± 0.2	1.3 ± 0.2
	12th	1.4 ± 0.3	1.4 ± 0.2	1.4 ± 0.3	1.4 ± 0.3	1.4 ± 0.3	1.3 ± 0.3	1.5 ± 0.2
LDL Cholesterol (mmol/l)	**B**	2.9 ± 0.8	2.8 ± 0.9	3.0 ± 0.7	2.8 ± 1.1	2.9 ± 0.6	2.8 ± 0.9	2.9 ± 0.8
	6th	2.3 ± 0.7	2.2 ± 0.7	2.4 ± 0.6	2.0 ± 0.7	2.4 ± 0.6	2.3 ± 0.7	2.3 ± 0.6
	12th	2.3 ± 0.8	2.2 ± 0.8	2.3 ± 0.8	2.1 ± 1.1	2.3 ± 0.6	2.2 ± 0.8	2.3 ± 0.7
Triglyceride (mmol/l)	**B**	1.4 ± 0.7	1.4 ± 0.6	1.4 ± 0.8	1.3 ± 0.6	1.5 ± 0.7	1.7 ± 0.9	1.3 ± 0.5
	6th	1.1 ± 0.4	1.1 ± 0.4	1.1 ± 0.3	0.9 ± 0.3	1.2 ± 0.3	1.2 ± 0.4	1.0 ± 0.3
	12th	1.0 ± 0.3	0.9 ± 0.3	1.0 ± 0.3	0.9 ± 0.3	1.0 ± 0.3	1.1 ± 0.3	0.9 ± 0.3
Fasting Glucose (mmol/l)	**B**	6.1 ± 1.0	5.9 ± 0.9	6.3 ± 1.2	6.1 ± 0.9	6.1 ± 1.1	6.8 ± 1.1	5.6 ± 0.6
	6th	5.4 ± 0.7	5.4 ± 0.6	5.4 ± 0.8	5.7 ± 0.8	5.3 ± 0.6	5.6 ± 1.0	5.3 ± 0.4
	12th	5.4 ± 0.9	5.4 ± 0.8	5.5 ± 1.0	5.8 ± 1.2	5.3 ± 0.7	5.7 ± 1.2	5.2 ± 0.5
Fasting Insulin (mU/l)	**B**	20.3 ± 17.2	20.5 ± 14.7	20.1 ± 20.0	19.1 ± 14.5	20.9 ± 18.5	24.8 ± 23.6	17.2 ± 9.8
	6th	8.2 ± 4.4	8.8 ± 4.6	7.7 ± 4.2	8.4 ± 4.6	8.2 ± 4.4	7.9 ± 3.9	8.5 ± 4.8
	12th	6.6 ± 4.1	7.1 ± 4.6	6.2 ± 3.5	6.7 ± 3.7	6.7 ± 4.3	7.0 ± 4.1	6.4 ± 4.1
Fat mass (kg)	**B**	64 ± 13	62 ± 12	65 ± 15	61 ± 16	65 ± 12	64 ± 12	63 ± 14
	6th	43 ± 12	41 ± 11	45 ± 12	39 ± 14	45 ± 10	39 ± 9	46 ± 13
	12th	38 ± 12	36 ± 13	39 ± 12	31 ± 14	40 ± 11	35 ± 11	40 ± 13
Fat ratio (percent)	**B**	49 ± 6	49 ± 6	49 ± 7	44 ± 7	51 ± 4	49 ± 5	49 ± 7
	6th	41 ± 7	41 ± 8	41 ± 7	35 ± 8	43 ± 5	39 ± 7	42 ± 8
	12th	37 ± 9	37 ± 10	38 ± 9	30 ± 9	41 ± 6	36 ± 8	39 ± 9

H Model Predictions vs. Data Distribution

Table H.1 | The comparison of HMOGP predictions (model) with fundamental statistical analysis on the data distribution (data). The comparison shows that fundamental statistical analysis of the data distribution gives unreliable peaking time (PT) and peak value (PV) results as they only rely on the timing of the observations. On the other hand, HMOGP predicts consistent AUC and TiH estimations with improved uncertainty compared to the conventional data analysis.

Glucose Response			**Baseline**	$p(\text{B}\rightarrow 6)$	**6th month**	$p(\text{B}\rightarrow 12)$	**12th month**	***Unit***
Peaking Time (PT)	**Data Statistics**	**OGTT**	60 (30-120)	0.409	30 (30-60)	0.417	30 (0-60)	***minute***
		p	0.250	-	0.391	-	0.386	
		MMT	60 (30-180)	0.150	30 (30-60)	0.179	30 (30-60)	
	HMOGP Predictions	**OGTT**	50 (45-62)	0.013	41 (35-45)	0.001	38 (33-42)	
		p	0.199	-	0.027	-	0.038	
		MMT	77 (44-129)	0.001	32 (28-37)	<0.001	31 (29-35)	
Peak Value (PV)	**Data Statistics**	**OGTT**	10.2 (6.2-16.5)	0.364	9.8 (6.5-17.3)	0.403	9.6 (6.3-21.7)	***mmol/l***
		p	0.105	-	0.297	-	0.371	
		MMT	8.0 (5.7-12.7)	0.162	11.1 (7.4-16.5)	0.205	10.7 (6.9-16.4)	
	HMOGP Predictions	**OGTT**	9.9 (9.4-10.4)	0.011	9.0 (8.5-9.4)	0.018	9.0 (8.5-9.5)	
		p	<0.001	-	<0.001	-	0.004	
		MMT	7.8 (7.4-8.2)	<0.001	10.3 (9.9-10.7)	<0.001	10.2 (9.7-10.6)	
Area Under Curve (AUC)	**Data Statistics**	**OGTT**	7.4 (5.3-13.5)	0.061	6.2 (4.9-11.9)	0.222	6.1 (4.7-13.7)	***mmol/l***
		p	0.368	-	0.312	-	0.371	
		MMT	7.1 (5.6-11.1)	0.300	6.9 (5.3-10.7)	0.308	6.3 (4.9-11.3)	
	HMOGP Predictions	**OGTT**	7.8 (7.4-8.2)	<0.001	6.1 (5.7-6.4)	<0.001	6.0 (5.6-6.4)	
		p	0.098	-	0.026	-	0.117	
		MMT	7.4 (7.0-7.7)	0.033	6.8 (6.4-7.1)	0.004	6.4 (6.0-6.9)	
Time in the Risk Zone (TiH)	**Data Statistics**	**OGTT**	104 (0-181)	0.136	70 (0-179)	0.236	60 (0-181)	***minute***
		p	0.474	-	0.391	-	0.414	
		MMT	120 (0-181)	0.400	78 (10-172)	0.397	62 (0-176)	
	HMOGP Predictions	**OGTT**	116 (103-135)	<0.001	66 (56-78)	<0.001	63 (54-75)	
		p	0.118	-	0.259	-	0.400	
		MMT	144 (109-166)	0.006	72 (63-93)	0.004	61 (55-78)	