



## OPEN ACCESS

EDITED BY  
Cecilia Diniz Behn,  
Colorado School of Mines, United States

REVIEWED BY  
Justin R. Ryder,  
University of Minnesota Twin Cities,  
United States  
Rongjing Ding,  
Peking University, China  
Sushant Singh,  
Amity University, Raipur, India

\*CORRESPONDENCE  
George Hripcsak,  
hripcsak@columbia.edu

SPECIALTY SECTION  
This article was submitted  
to Metabolic Physiology,  
a section of the journal  
Frontiers in Physiology

RECEIVED 19 April 2022  
ACCEPTED 11 October 2022  
PUBLISHED 28 November 2022

CITATION  
Richter LR, Albert BI, Zhang L,  
Ostropolets A, Zitsman JL, Fennoy I,  
Albers DJ and Hripcsak G (2022), Data  
assimilation on mechanistic models of  
glucose metabolism predicts glycemic  
states in adolescents following  
bariatric surgery.  
*Front. Physiol.* 13:923704.  
doi: 10.3389/fphys.2022.923704

COPYRIGHT  
© 2022 Richter, Albert, Zhang,  
Ostropolets, Zitsman, Fennoy, Albers  
and Hripcsak. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Data assimilation on mechanistic models of glucose metabolism predicts glycemic states in adolescents following bariatric surgery

Lauren R. Richter<sup>1</sup>, Benjamin I. Albert<sup>1</sup>, Linying Zhang<sup>1</sup>,  
Anna Ostropolets<sup>1</sup>, Jeffrey L. Zitsman<sup>2</sup>, Ilene Fennoy<sup>3</sup>,  
David J. Albers<sup>1,4,5</sup> and George Hripcsak<sup>1\*</sup>

<sup>1</sup>Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY, United States, <sup>2</sup>Division of Pediatric Surgery, Department of Surgery, Columbia University Irving Medical Center, New York, NY, United States, <sup>3</sup>Division of Pediatric Endocrinology, Metabolism, and Diabetes, Department of Pediatrics, Columbia University Irving Medical Center, New York, NY, United States, <sup>4</sup>Department of Bioengineering, University of Colorado Anschutz Medical Campus, Aurora, CO, United States, <sup>5</sup>Department of Biomedical Informatics, University of Colorado Anschutz Medical Campus, Aurora, CO, United States

Type 2 diabetes mellitus is a complex and under-treated disorder closely intertwined with obesity. Adolescents with severe obesity and type 2 diabetes have a more aggressive disease compared to adults, with a rapid decline in pancreatic  $\beta$  cell function and increased incidence of comorbidities. Given the relative paucity of pharmacotherapies, bariatric surgery has become increasingly used as a therapeutic option. However, subsets of this population have sub-optimal outcomes with either inadequate weight loss or little improvement in disease. Predicting which patients will benefit from surgery is a difficult task and detailed physiological characteristics of patients who do not respond to treatment are generally unknown. Identifying physiological predictors of surgical response therefore has the potential to reveal both novel phenotypes of disease as well as therapeutic targets. We leverage data assimilation paired with mechanistic models of glucose metabolism to estimate pre-operative physiological states of bariatric surgery patients, thereby identifying latent phenotypes of impaired glucose metabolism. Specifically, maximal insulin secretion capacity,  $\sigma$ , and insulin sensitivity,  $S_I$ , differentiate aberrations in glucose metabolism underlying an individual's disease. Using multivariable logistic regression, we combine clinical data with data assimilation to predict post-operative glycemic outcomes at 12 months. Models using data assimilation sans insulin had comparable performance to models using oral glucose tolerance test glucose and insulin. Our best performing models used data assimilation and had an area under the receiver operating characteristic curve of 0.77 (95% confidence interval 0.7665, 0.7734) and mean average precision of 0.6258 (0.6206, 0.6311). We show that data assimilation extracts knowledge from mechanistic models of glucose metabolism to infer future glycemic states from limited clinical data. This method can provide a pathway to predict long-term, post-surgical glycemic states by estimating the

contributions of insulin resistance and limitations of insulin secretion to pre-operative glucose metabolism.

#### KEYWORDS

type 2 diabetes, data assimilation, mechanistic models of glucose metabolism, pediatrics, bariatric surgery, machine learning, obesity

## 1 Introduction

As obesity rates rise in the United States, so too does the prevalence of type 2 diabetes mellitus (T2DM) in children and adolescents (Skelton et al., 2009; Hales et al., 2017h). While a number of pharmacotherapies exist to treat T2DM, there are few options approved for use in younger patients, who typically have more aggressive disease (TODAY Study Group, 2021; American Diabetes Association Professional Practice Committee, 2022). As such, bariatric surgery is increasingly used as a treatment for severe obesity and prevention or reversal of T2DM, despite risk of operative complications (Hsia et al., 2012; Aung et al., 2016; Beamish and Reinehr, 2017; Inge et al., 2018a; Armstrong et al., 2019; Bolling et al., 2019; Karasko, 2019; Khattab and Sperling, 2019; American Diabetes Association Professional Practice Committee Draznin et al., 2022). Many patients benefit with significant, sustained weight loss, improvement in quality of life, and improvement of obesity-related comorbidities (Inge et al., 2016; Rubino et al., 2016; Beamish and Reinehr, 2017; Pedroso et al., 2018; Inge et al., 2019). However, an ill-defined subset of this population have sub-optimal outcomes (Montero et al., 2011; Livhits et al., 2012; Toh et al., 2017; Inge et al., 2018a; Inge et al., 2018b). Predicting which patients are most likely to benefit from surgery and how they will benefit is a current challenge aimed to minimize unnecessary risk during a critical period of growth and development.

The prevalence of impaired glucose metabolism (IGM)—here referring to a heterogenous population with impaired glucose tolerance (IGT), impaired fasting glucose (IFG), prediabetes (preDM), or T2DM— is increasing in pediatric populations, and may be underestimated (Sinha et al., 2002; Lee et al., 2006; Nowicka et al., 2011; Buse et al., 2013; Dabelea et al., 2014). IGM is progressive and often goes undiagnosed until later in disease history. It is associated with insulin resistance (IR), the need for more insulin to achieve physiologic effects, i.e., peripheral glucose uptake and suppression of hepatic glucose production (HGP). In the obese state, IR is almost guaranteed as it is directly related to visceral adiposity (Kahn and Flier, 2000), but the extent to which the pancreas can compensate exists on a spectrum. Insulin sensitivity ( $S_I$ ) is a measure of the effectiveness of insulin in promoting glucose uptake. It is reciprocally related to insulin resistance. In this context, T2DM is an example

extreme IGM, with significant IR coupled with  $\beta$  cell dysfunction and resultant hyperglycemia (Prentki and Nolan, 2006; American Diabetes Association Professional Practice Committee, 2022). The extent to which function can be rescued after progression to T2DM is modifiable to some degree (Lim et al., 2011; Pajvani and Accili, 2015; Taylor et al., 2019; Richter et al., 2020; Holst and Madsbad, 2021; Bartolomé et al., 2022). Duration of disease negatively impacts the probability of resolution, and lifestyle interventions resulting in weight loss during the early prediabetic phase are more likely to prevent progression (Knowler et al., 2002).

Reversal of preDM and prevention of T2DM are therefore considered to be one of the major benefits of bariatric surgery in this age group for which there are limited other options (Aung et al., 2016; Armstrong et al., 2019; Bolling et al., 2019; Khattab and Sperling, 2019; American Diabetes Association Professional Practice Committee Draznin et al., 2022). Indeed, when compared to pharmacologic or lifestyle interventions, bariatric surgery is overall the most successful intervention with respect to glycemic improvements and sustained weight loss (Schauer et al., 2012; Courcoulas et al., 2014; Mingrone et al., 2015). The specific surgery has clear impact on outcomes. The three most prevalent bariatric surgeries in the U.S. are adjustable gastric banding (AGB), Roux-en-Y gastric bypass (RYGB), and vertical sleeve gastrectomy (VSG). In AGB, an inflatable band is placed around the upper part of the stomach creating a small pouch. In RYGB, the jejunum is directly connected to a remnant small pouch of stomach, thereby bypassing the majority of the stomach and the duodenum. In VSG, the majority of the stomach is removed along the greater curvature, creating a narrow tube or sleeve. While AGB is thought to act through purely restrictive mechanisms, VSG and RYGB restrict food intake and increase malabsorption, alter secretion of gut hormones related to satiety and insulin secretion, e.g., glucagon-like peptide-1 (GLP-1) and ghrelin, and change bile acid composition through the change in macronutrients present in areas of the small intestine (Seeley et al., 2015; Mulla et al., 2018; Akalestou et al., 2022). These differing effects can have profound impact on insulin resistance in particular, and the success of VSG and RYGB in comparison to AGB has led to their increased usage (Mulla et al., 2018). Depending on how remission is defined, meta-analyses have shown that 20–80% of adults will have some degree of improvement in T2DM at medium-to-long term follow up (Yip et al., 2013; Elbahrawy et al., 2018; Tsilingiris et al., 2019; Purnell et al., 2021),

although there is suggestion of continued impaired  $\beta$  cell function and overestimation of success (Ramos-Levi et al., 2013a; Laferrère and Pattou, 2018). Smaller prospective studies focusing on adolescents suggest T2DM remission may occur in up to 80–90% of patients and improvement of preDM may occur in 70–80% of patients following Roux-en-Y gastric bypass (RYGB), the most drastic surgery that is recommended for adolescents with respect to metabolic intervention and risk of complications (Inge et al., 2014; Inge et al., 2017; Olbers et al., 2017; Stefater and Inge, 2017). However, small sample sizes and relatively homogenous populations (> 70% non-Hispanic white) limit generalizability of results. Predicting which patients are likely to have remission or partial remission of IGM as a result of surgery and thus not develop T2DM remains a challenging task (Ramos-Levi et al., 2013b; Wang et al., 2015; Tsilingiris et al., 2019).

Given the complexity of medical decision making in adolescent bariatric surgery, accurately assessing benefits vs. risks for an individual is critical for patients and their care teams. Glucose dysregulation and obesity can compound existing surgical risks and increase the chance of complications. Although adolescents have lower complication rates compared to their adult counterparts, they may still experience wound infections, anastomotic strictures, leaks, wound dehiscence, abdominal hernias, dehydration, and venous thromboembolism (Lamoshi et al., 2020). In patients with diabetes, poor wound healing and risk of infection are serious considerations and can directly impact the success of the surgery and need for revision (Keidar, 2011). In the long term, patients also have significant malabsorption post-bariatric surgery, resulting in multiple vitamin and mineral deficiencies requiring lifelong supplementation (Bal et al., 2012; Lamoshi et al., 2020). The consequences of this decreased nutrition can be great in adolescents who are still undergoing their growth spurts and accruing bone during their pubertal years (Lamoshi et al., 2020; Weiner et al., 2020; Ou et al., 2022). Although advancements have been made to minimize risk of post-operative complications, improve wound healing, and optimize post-operative nutrition, bariatric surgery remains an aggressive measure taken to improve a patient's health (Crossan and Sheer, 2022; Dewberry et al., 2022). Therefore, providing accurate information about a patient's current diabetic state the probability of improvement with surgery can allow for more informed decision making.

Previous studies using statistical or traditional machine learning techniques to predict T2DM outcomes have generally focused on relatively homogenous adult populations who underwent RYGB, with or without genetic information included in analysis. Because available data are sparse, these approaches are subject to error and are not often validated in pediatric populations (Aminian et al., 2017; Cao et al., 2020; Kam et al., 2020). Varying definitions of

what successful glycemic outcomes mean also complicate prediction (Buse et al., 2009; Holst and Madsbad, 2021). Outside of surgery type, other potential predictors include anthropometrics (weight, height, body mass index [BMI]), pre-operative disease severity, use of anti-diabetic medications, and presence of comorbidities (DeMaria et al., 2007; Livhits et al., 2012; Dixon et al., 2013; Robert et al., 2013; Panunzi et al., 2015; Wang et al., 2015; Shen et al., 2019). Associations are also seen with baseline biomarkers such as fasting glucose, insulin, C-peptide, triglycerides (TG), C-reactive protein (CRP), and hemoglobin A1c (HbA1c) levels (Ortega et al., 2012; Courcoulas et al., 2013; Inge et al., 2014; Pedersen et al., 2016; Yan et al., 2017). These features, particularly pre-operative disease severity (as measured by duration of disease, labs, and medications) are more homogenous in adolescents, who frequently have preDM rather than T2DM, and are therefore on fewer medications, if any. Additionally, adolescents tend to have lower HbA1c values at baseline prior to undergoing bariatric surgery compared to their adult counterparts. As such, features that are useful predictors in adult surgical response are not necessarily translatable to adolescents.

To provide more personalized predictions, Pedersen, et al. incorporated genetic and clinical information in an artificial neural network to accurately predict short-term discontinuation of diabetes medications at 30 days (Pedersen et al., 2016). The majority of candidate genetic markers were associated with insulin secretion, glucose clearance, or insulin sensitization. While genetics certainly play a role in T2DM, pre-operative genetic analyses are not currently practical for every patient (Hatoum et al., 2011; Okser et al., 2013; Rouskas et al., 2014).

These prior studies provide evidence that an individual's underlying physiology has long-term implications for treatment outcomes, but existing methods to approximate these physiological pathways may not be practical for use in adolescents (Lee, 2007; Brown and Yanovski, 2014). Glucose tolerance and insulin resistance are frequently estimated using point or dynamic lab proxies due to the expensive and invasive nature of the gold standard for measurement, the hyperinsulinemic-euglycemic clamp (Muniyappa et al., 2018). Oral glucose tolerance tests (OGTTs) are one such approximation used to screen for and diagnose dysglycemia (Olson et al., 2010; Muniyappa et al., 2018; American Diabetes Association Professional Practice Committee, 2022). After fasting (usually overnight), patients are given a fixed dose of liquid glucose (typically 75 g) after which glucose levels are measured at timed intervals. OGTTs can vary in the types of labs drawn, the frequency of sampling, and the duration of the procedure (Muniyappa et al., 2018). In common clinical practice, glucose is measured over two or three hours. Measurements of insulin and C-peptide are not the standard of care; their assays are

relatively costly and lack of standardization makes their interpretation less straightforward (Manley et al., 2007; Little et al., 2008; Miller et al., 2009; Tohidi et al., 2017). As such, these labs are typically only collected in research settings.

Fasting insulin and glucose measurements can be used to calculate indices such as the homeostatic model assessment of insulin resistance (HOMA-IR), which can approximate parameters such as peripheral and hepatic insulin sensitivity  $hepa_{S_i}$  as would be assessed in clamp studies or frequently sampled intravenous GTTs (FSIVGTT), a silver standard (Matthews et al., 1985; Bergman et al., 1987; Yeckel et al., 2004). Interpretation of these indices can vary by patient characteristics (e.g., age, sex, ethnicity, body habitus) (Wallace et al., 2004; Gunczler and Lanes, 2006; Nathan et al., 2007; Shaibi et al., 2011; Gutch et al., 2015; Arslanian et al., 2019; Tagi et al., 2019; Kim et al., 2020). In particular, HOMA-IR may not be sensitive to improvements in insulin sensitivity in adolescents (Shaibi et al., 2011; Bryant et al., 2014).

Several mechanistic models of glucose and insulin metabolism have been empirically developed and validated against clamp or FSIVGTT studies to mathematically describe glucose metabolism. Models such as the meal model (Dalla Man et al., 2007), oral minimal model (Cobelli et al., 2014), and ultradian model (Sturis et al., 1991) incorporate insulin secretion rates and glucose elimination to estimate states over varying timescales. Topp et al. developed a mechanistic model that incorporates more granular  $\beta$  cell dynamics, which was extended by Ha et al. to model the development of type 2 diabetes over time and quantify the extent to which different insults in the system contribute to disease (Topp et al., 2000; Ha et al., 2016; Ha and Sherman, 2020). Represented as a system of ordinary differential equations (ODEs), these models have the potential to allow for patient-level characterization of glucose metabolism.

Extracting clinical knowledge from these models is not straightforward, as both the models themselves and their results are often viewed as too abstract for application in clinical practice. However, within these mechanistic models are clinically meaningful physiologic parameters when applied to the appropriate problems, e.g. assessing insulin sensitivity's relationship with lipoprotein metabolism (Chung et al., 2022). Furthermore, while clamp studies represent a patient's physiologic state and glucose excursions at a specific point in time, mechanistic models have the potential to elucidate more long-term physiologic states that are difficult to capture clinically.

Knowledge within mechanistic models of glucose metabolism can be exploited via data assimilation, a family of methods frequently used in meteorology and aerospace science (Evensen, 2009; Law et al., 2015). Data assimilation leverages the underlying information about the system contained in these mechanistic models to update current and past states using filtering and smoothing, updates that in turn provide the ability to forecast future states by running the estimated

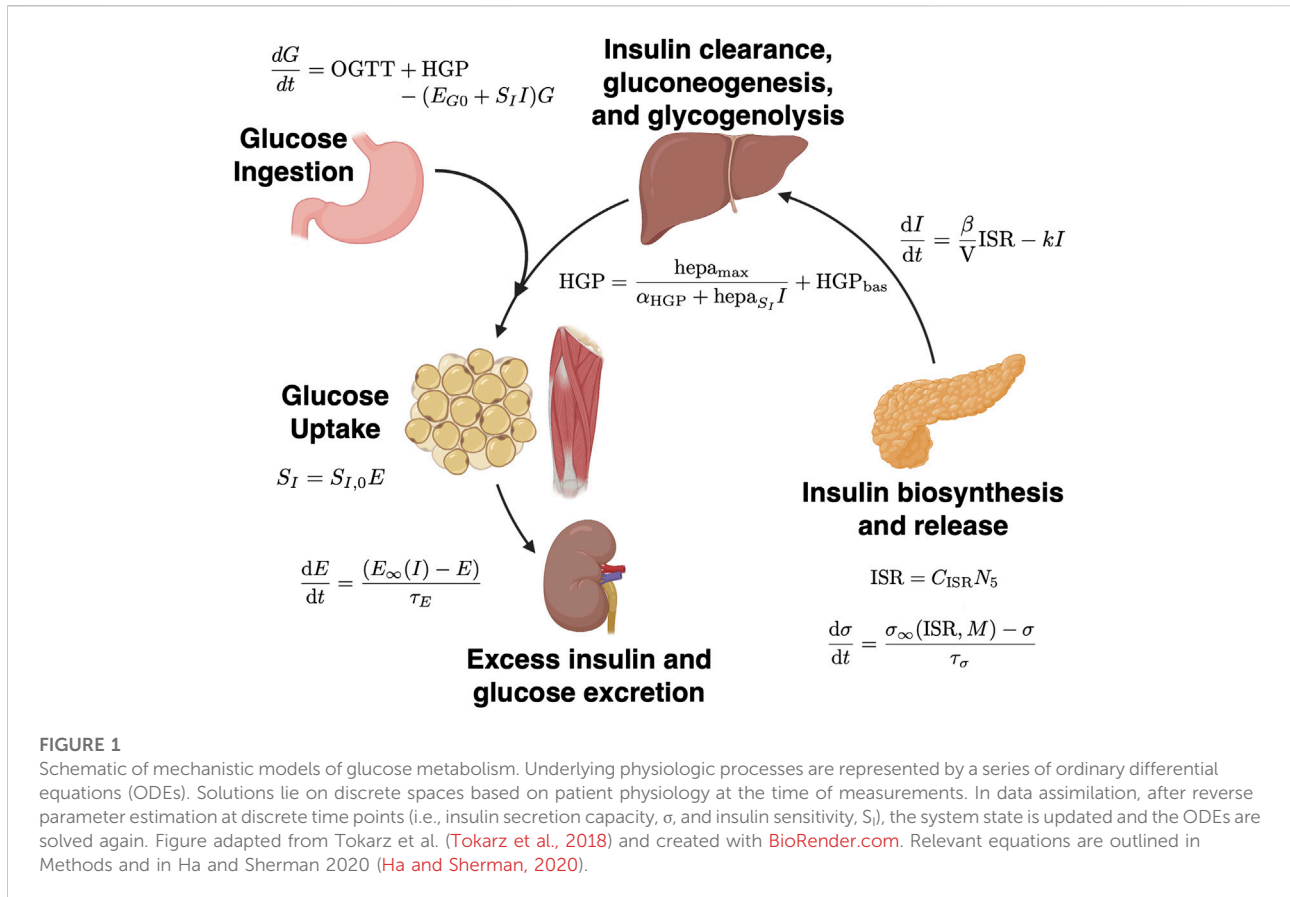
model forward in time. Various filters exist that allow for parameterization and propagation of state uncertainty for non-linear systems such as glucose homeostasis (Kalman, 1960; Wan et al., 2001; Julier and Uhlmann, 2004). In previous work, we used the ultradian and meal models with an unscented Kalman filter (UKF), a sequential method that propagates uncertainty in non-linear systems, as well as deterministic and stochastic optimization methods using techniques such as Markov chain Monte Carlo (MCMC) to estimate both states and parameters, generating a real-time, personalized forecast from free-living data (Albers et al., 2017; Levine et al., 2017). In the free-living data context where data are sparse, we developed a constrained (Albers et al., 2019) version of the ensemble Kalman filter (EnKF) method (Evensen, 2009) that made state and parameter estimates more robust. We also successfully applied these methods to real-time glucose forecasting in the setting of T2DM with OGTT data (Mulgrave et al., 2020). Whereas in the free-living case the focus was accurate glucose state determination, the OGTT case focused more on parameter estimates as a marker for underlying disease.

Here, we take this prior work to further demonstrate the validity of using physiologic parameters inferred from mechanistic models of glucose metabolism to predict impaired glucose metabolism (IGM) in adolescents at 12-months post-bariatric surgery as compared to other clinical information. We use data assimilation to estimate parameters from an extension of the model initially proposed by Topp et al. incorporating  $\beta$  cell mass (Ha and Sherman, 2020), partially represented in Figure 1. We then train logistic regression models leveraging these parameter estimates derived from data assimilation on a cohort of adolescents who underwent vertical sleeve gastrectomy (VSG) or laparoscopic adjustable gastric banding (AGB) at our institution between 2006 and 2020.

## 2 Materials and methods

### 2.1 Extraction of clinical data

Data were collected retrospectively from Columbia University Irving Medical Center (CUIMC) from adolescent patients aged 10–21 who had bariatric surgery between 2006 and 2020. Records were first selected based on the presence of bariatric surgery procedure codes with a diagnosis of obesity on the same day ( $n = 396$ ), a 120-minute OGTT measuring glucose and insulin (at 0, 30, 60, and 120 min) within one year prior to surgery ( $n = 202$ ), a pre-operative HbA1c within 120 days of the OGTT ( $n = 202$ ), and at least one post-operative outcome documented within 6–18 months post-surgery ( $n = 176$ ). These patients were seen through the Center for Adolescent Bariatric Surgery (CABS) at CUIMC. Patients with diagnosis codes associated with type 1, cystic fibrosis-related, or



gestational diabetes mellitus were excluded. Features as outlined above were extracted for patients who met criteria.

## 2.2 Data pre-processing and manual feature selection

In addition to features necessary to include patients (OGTT glucose and insulin measurements and HbA1c), additional laboratory variables were selected *a priori*. These labs were pre-operative thyroid stimulating hormone (TSH), thyroxine (T4), free T4 (FT4), triiodothyronine (T3), aspartate transaminase (AST), alanine transaminase (ALT), total cholesterol (TC), high-density lipoprotein (HDL), low-density lipoprotein (LDL), triglycerides (TG), and C-peptide. Age was used as a continuous variable. Categorical features included surgery type, demographic data (self-reported race, ethnicity, and sex), and pre-operative presence of specific comorbidities associated with insulin resistance and T2DM: liver disease (including non-alcoholic steatohepatitis [NASH] and non-alcoholic fatty liver disease [NAFLD]), polycystic ovary syndrome (PCOS), dyslipidemia, hypertension (HTN),

obstructive sleep apnea (OSA), thyroid disease, metabolic syndrome (MetS), T2DM (by diagnosis code), and abnormal glucose metabolism (by diagnosis code). To be coded as having the comorbidity of interest without manual chart review, at least 25% of all encounters within the appropriate time window had to contain a related diagnosis code (5 years pre-operative; 6–18 months post-operative) (Perotte and Hripcsak, 2013). Features missing in more than 25% of the sample were removed from analysis. The closest labs obtained prior to surgery and the pre-operative OGTT were included as features if there were multiple lab results within the appropriate time window. HOMA-IR was calculated using mass units (Matthews et al., 1985) as follows in Eq. 1:

$$\text{IR}_{\text{HOMA}} = \frac{G_0 \times I_0}{405} \tag{1}$$

where  $G_0$  is fasting glucose in mg/dL and  $I_0$  is fasting insulin in  $\mu\text{IU/mL}$ .

For continuous variables, outliers were defined as those values outside  $1.5\times$  the interquartile range (IQR). However, after manual chart review, none of the outliers were removed as none were found to be spurious measurements.

**TABLE 1** Outcome definitions for classification as having impaired glucose metabolism (IGM) or normal glucose metabolism (NGM) at 12 months post-bariatric surgery.

	Normal glucose metabolism (NGM)	Impaired glucose metabolism (IGM)	
		Prediabetes	Type 2 diabetes mellitus
HbA1c (%)	< 5.7	5.7–6.4	≥ 6.5
G <sub>0</sub> (mg/dL)	< 100	100–125	≥ 126
G <sub>120</sub> (mg/dL)	< 140	140–199	≥ 200
Anti-diabetic medications	None	Metformin or GLP-1a	All other drug classes
Diagnosis codes in encounters within 6–18 months post-op	None or < 25%	Abnormal Glucose in ≥ 25% of encounters	T2DM in ≥ 25% of encounters

G, glucose; GLP-1a, Glucagon-Like Peptide-1 agonist; OGTT, oral glucose tolerance test; OSA, obstructive sleep apnea; T2DM, type 2 diabetes mellitus.

Due to non-normal distributions of the majority of continuous variables, numeric features were scaled and standardized using power transformation with the Box-Cox method (Box and Cox, 1964) prior to hyperparameter tuning. Missing values ( $n = 7$ , all thyroid function tests) were imputed using iterative imputation using five nearest features with sampling from the prior distribution (Buck, 1960; Buuren and Groothuis-Oudshoorn, 2011; Pedregosa et al., 2011).

In total, we used 21 continuous clinical variables, 23 categorical variables (three multinomial, 11 binary), and six continuous parameter estimates from data assimilation (mean and standard deviations of maximal insulin secretion capacity [ $\sigma$ ], insulin sensitivity [ $S_I$ ], and their product [ $\sigma^*S_I$ ] for an individual's estimated distributions). Models including HOMA-IR as a feature did not include data assimilation estimates. After data pre-processing, a total of 49 features were included in the most comprehensive model.

## 2.3 Outcome labelling

Patient post-operative outcome classification was coded as a binary variable indicating the presence or absence of impaired glucose metabolism (IGM) at 12 months. To be classified as having IGM, patients could have any one of the following criteria occur in the 6–18-month post-operative window: ≥ 2 elevated HbA1c values, post-operative OGTT  $G_0 \geq 100$  mg/dL, post-operative OGTT  $G_{120} \geq 140$  mg/dL, anti-diabetic drug use (including metformin), or presence of diagnosis codes for T2DM or abnormal glucose metabolism in ≥ 25% of encounters in the post-op window (American Diabetes Association Professional Practice Committee, 2022). If multiple specimens for the same lab were collected within the 6–18-month window, the latest labs were used. Outcome definitions are shown in Table 1.

## 2.4 Predictive model training and evaluation

Regularized logistic regression models were trained to predict impaired glucose metabolism (IGM) as a binary outcome on varied subsets of the features as input. The data were first split into 70–30 train-test sets. Hyperparameters were tuned on the train set with stratified nested  $k$ -fold cross-validation. Tuned hyperparameters included regularization method ( $L2$  vs.  $L1$ ), regularization constant ( $\lambda = \alpha^*n$ ), learning rate [ $f(\alpha)$ ], and max iterations. The loss function was binary cross-entropy loss with balanced class weight. The optimizer was stochastic gradient descent with an adaptive learning rate. To ensure model robustness to random data splitting, we performed 30 independent train-test splits and the results were averaged. The hyperparameters associated with the minimum loss in training were selected for model evaluation.

Because our dataset was imbalanced with respect to outcomes, we chose class weights in the logistic regression models that were inversely proportional to the class frequency. To evaluate the performance of the prediction models, we computed area under the Receiver Operating Characteristic curve (AUROC), precision, recall, and average precision (AP) on out-of-bag estimates from 1,000 bootstrapped samples ( $n = 176$ ). Average precision refers to the weighted mean of precision with respect to recall at all probability thresholds for classification and focuses on how well the models predict the positive class (in our case, post-operative IGM). It is analogous to the area under the precision-recall curve (AUPRC) and better suited for imbalanced datasets compared to accuracy. Baseline performance of a naïve classifier would have an AUPRC equal to the proportion of the positive class in the data (here, 0.318).

## 2.5 Mechanistic glucose metabolism modeling

Relevant equations used in our study are briefly outlined below, with a full description of the model available in the supplemental materials of Ha and Sherman 2020 (Ha and Sherman, 2020; Sherman, 2022).

Eq. 2 describes change in glucose over time as a function of the glucose flux during the OGTT (OGTT), hepatic glucose production (HGP), insulin sensitivity ( $S_I$ ), the insulin-independent effectiveness of glucose ( $E_{G0}$ ), current glucose ( $G$ ), and insulin ( $I$ ). It is given as:

$$\frac{dG}{dt} = \text{OGTT} + \text{HGP} - (E_{G0} + S_I I)G \quad (2)$$

where HGP is a decreasing function of  $I$  and hepatic insulin sensitivity (hepa $s_I$ ) and  $E_{G0}$  represents insulin-independent glucose uptake by peripheral tissues, here fixed at  $0.0118 \text{ min}^{-1}$ .

Eqs. 3, 4 describe change in insulin over time as a function of  $\beta$  cell mass ( $\beta$ ), volume of distribution ( $V$ ), insulin secretion rate (ISR), and insulin ( $I$ ). ISR is a function of calcium ion concentration in the  $\beta$  cell cytosol ( $C_{\text{ISR}}$ ) and the number of primed insulin vesicles at the  $\beta$  cell membrane ( $N_5$ ). They are given as:

$$\frac{dI}{dt} = \frac{\beta}{V} \text{ISR} - kI \quad (3)$$

$$\text{ISR} = C_{\text{ISR}} N_5 \quad (4)$$

where  $k$  is a rate constant of insulin clearance. The precise form of  $C_{\text{ISR}}$  as a function of glucose ( $G$ ) and ATP-dependent potassium ion channel ( $\text{K}^+$ -ATP) density ( $\gamma$ ) can be determined by expanding the equations derived from a slight modification of the steady state of the previously published exocytosis model (Chen et al., 2008).

Eq. 5 shows the change in  $\gamma$ , the density of the  $\beta$  cell membrane ATP-dependent potassium ion channel ( $\text{K}^+$ -ATP), as an increasing function of glucose ( $G$ ) on a scale of hours to days. This describes the shift in glucose-dependent insulin secretion in the setting of chronic hyperglycemia, where increases in channel density lead to decreases in insulin secretion. Because of the short time scale of the OGTT,  $\gamma$  was fixed at  $-0.076$  as in prior work (Sherman, 2022) and this equation was not explicitly solved in our methods. It is provided here for clarity on how chronic hyperglycemia affects  $\beta$  cell structure and is given as:

$$\frac{d\gamma}{dt} = \frac{\gamma_{\infty}(G) - \gamma}{\tau_{\gamma}} \quad (5)$$

where  $\gamma_{\infty}(G)$  is an increasing sigmoidal function of glucose and  $\tau_{\gamma}$  is a time constant.

From the insulin exocytosis model, Eqs. 6, 7 describe the change in the number of vesicles in the  $\beta$  cell granule-membrane

complex during docking and priming ( $N_6$  and  $N_5$ , respectively) as functions of  $\text{K}^+$ -ATP channel density ( $\gamma$ ), glucose ( $G$ ), maximal insulin secretion capacity ( $\sigma$ ), and the baseline insulin vesicle priming rate,  $r_2^0$  (Chen et al., 2008). These equations are given as:

$$\frac{dN_5}{dt} = C_{5,5} N_5 + C_{5,6} N_6 \quad (6)$$

$$\frac{dN_6}{dt} = C_{6,0} + C_{6,5} N_5 + C_{6,6} N_6 \quad (7)$$

where  $C_{i,j}$  represents the cytosol calcium ion concentration at a given state in exocytosis:  $C_{5,5}$  is a function of  $G$  and  $\gamma$ ;  $C_{5,6}$  and  $C_{6,6}$  are functions of  $G$ ,  $\gamma$ , and  $r_2^0$ ;  $C_{6,0}$  is a function of  $G$ ,  $\gamma$ , and  $\sigma$ , and;  $C_{6,5}$  is a constant.

Eq. 8 describes the change in the maximal insulin secretion capacity,  $\sigma$ , to compensate for chronic hyperglycemia on a scale of hours to days. It is a unitless scale factor. During the OGTT,  $\sigma$  is assumed to be at steady state.  $\sigma$  is a function of insulin secretion rate (ISR) and  $\beta$  cell metabolism ( $M$ ), where  $\sigma$  increases with increases in ISR and decreases as  $M$  increases.

$$\frac{d\sigma}{dt} = \frac{\sigma_{\infty}(\text{ISR}, M) - \sigma}{\tau_{\sigma}} \quad (8)$$

Parameters and non-estimated initial states were set according to their values in previous research (Ha and Sherman, 2020; Sherman, 2022). In experiments where measured insulin values were not included in the estimation optimization,  $I_0$  was set as  $5.63 \mu\text{IU/mL}$ . The parameters estimated without using a patient's insulin values are denoted with an,  $\overline{\sigma S_I}$  and  $\overline{\sigma^* S_I}$ .

## 2.6 Estimating parameters using data assimilation

In our methods, we estimated two parameters via data assimilation, insulin sensitivity,  $S_I$ , and maximal insulin secretion capacity,  $\sigma$ . Ordinary differential equations were solved using a Rosenbrock- $W$  method (Rosenbrock23) (Rackauckas and Nie, 2017) within the bounds of  $[0.005, 3]$  for  $S_I$  and  $[0.01, 10]$  for  $\sigma$ . The posterior distributions of the parameters and the product  $\sigma^* S_I$  were estimated based on a 500,000 iteration Random Walk Metropolis-Hastings MCMC (Hastings, 1970) chain (excluding a burn-in of 50,000 iterations), assuming a normal distribution error model and uniform priors. A decrease in autocorrelation approaching 0 was appreciated with increasing lags ( $k$ ) near  $k = 500$  to support chain convergence, and the number of iterations and burn-in were selected to be orders of magnitude larger than  $k$  (Roy, 2019). Acceptance rate varied with each patient between 0.2–0.5. Multiple independent chains for a subset of 10 patients provided confidence in the reproducibility of the estimates. The means and standard deviations of the parameters were calculated from the remaining 450,000 iterations (Ge et al., 2018) in an individual's chain, and these summary statistics were included in logistic regression models.

**TABLE 2** Pre-operative demographic comparisons by binary post-operative outcome at 12 months. For continuous variables, medians with their 95% CI are shown. Categorical variables are shown as a percentage and count. Significant *p*-values for Mann-Whitney-U tests after Bonferroni correction are in bold (*p* < 0.001).

Variable Name	– IGM <i>n</i> = 120	+ IGM <i>n</i> = 56	<i>p</i> -value
<b>Pre-operative IGM status<sup>‡</sup></b>	<b>45 (54)</b>	<b>83.9 (47)</b>	<b>&lt; 0.0001</b>
Operation Age (years)	16.9 (16.6, 17.1)	16.8 (16.5, 17.5)	0.3034
Pre-operative BMI (kg/m <sup>2</sup> )	45.93 (44.66, 47.92)	48.25 (46.20, 51.24)	0.0712
Male (%)	27.5 (33)	33.9 (19)	0.4880
<b>Race</b>			
Asian	0 (0)	3.6 (2)	0.1000
Black	17.5 (21)	28.6 (16)	0.1390
Native American	0 (0)	1.8 (1)	0.3180
Other	2.5 (3)	7.1 (4)	0.2110
Pacific Islander	0.8 (1)	0 (0)	1.0000
White	58.3 (70)	46.4 (26)	0.1890
Unknown	20.8 (25)	16.1 (9)	0.5890
<b>Ethnicity</b>			
Hispanic	42.5 (51)	51.8 (29)	0.3220
Non-Hispanic	53.3 (64)	51.8 (29)	0.9760
Declined	1.7 (2)	1.8 (1)	1.0000
Unknown	9.2 (11)	1.8 (1)	0.1060
<b>Surgery Type</b>			
Roux-en-Y (RYGB)	0.8 (1)	0 (0)	1.0000
Gastric Band (AGB)	42.5 (51)	60.7 (34)	0.0370
Sleeve Gastrectomy (VSG)	55.8 (67)	39.3 (22)	0.0600
Other <sup>‡</sup>	0.8 (1)	0 (0)	1.0000
<b>ICD Codes</b>			
Abnormal Glucose <sup>‡</sup>	13.3 (16)	33.9 (19)	0.0030
T2DM	8.3 (10)	23.2 (13)	0.0130
Dyslipidemia	29.2 (35)	35.7 (20)	0.4850
GERD	13.3 (16)	14.3 (8)	1.0000
Hypertension	21.7 (26)	41.1 (23)	0.0130
Liver Disease <sup>‡</sup>	17.5 (21)	21.4 (12)	0.6780
Metabolic Syndrome	29.2 (35)	32.1 (18)	0.8220
OSA	35 (42)	42.9 (24)	0.4030
PCOS (F, <i>n</i> = 87, 37)	19.5 (17)	48.6 (18)	0.0021
Thyroid Disease	4.2 (5)	8.9 (5)	0.2930

<sup>‡</sup>Patients with pre-operative IGM had any one of the IGM definitions pre-operatively. This variable was not used in training any of the models.

<sup>‡</sup>Other surgeries refers to non-specific procedure codes for restrictive bariatric surgery.

<sup>‡</sup>Liver disease includes non-alcoholic fatty liver disease, non-alcoholic steatohepatitis, and unspecified chronic liver disease.

<sup>‡</sup>Abnormal glucose diagnosis codes exclude any kind of diabetes mellitus.

Abbreviations: BMI, body mass index; CI, confidence interval; GERD, gastroesophageal reflux disease; IGM, impaired glucose metabolism; OSA, obstructive sleep apnea; PCOS, polycystic ovary syndrome; T2DM, type 2 diabetes mellitus.

## 2.7 Descriptive statistics

The majority of continuous features were not normally distributed, so non-parametric methods were used to calculate

descriptive statistics on non-transformed data. These data are reported as medians with 95% confidence intervals around the median. The Mann-Whitney-U tests were used to compare continuous variables under these circumstances.  $\chi^2$  or Fisher's



TABLE 3 Pre-operative laboratory value comparisons by binary post-operative outcome at 12 months. Medians with their 95% CI are shown. Significant *p*-values for Mann-Whitney-U tests after Bonferroni correction are in bold (*p* < 0.002).

Variable Name	– NGM <i>n</i> = 120	+ IGM <i>n</i> = 56	<i>p</i> -value
Baseline HbA1c (%)	5.4 (5.3, 5.5)	5.8 (5.7, 6.0)	< 0.0001
Baseline HOMA-IR	3.13 (2.76, 3.52)	4.53 (3.75, 5.57)	1.987 × 10 <sup>-3</sup>
<b>OGTT Measurements</b>			
G <sub>0</sub> (mg/dL)	86.5 (84, 88)	87.5 (86, 93)	0.1081
G <sub>30</sub> (mg/dL)	128 (123, 134)	140.5 (134, 146)	1.312 × 10 <sup>-3</sup>
G <sub>60</sub> (mg/dL)	115.5 (113, 121)	133.5 (127, 154)	2.2 × 10 <sup>-4</sup>
G <sub>120</sub> (mg/dL)	98 (93, 104)	113.5 (109, 122)	1.5 × 10 <sup>-4</sup>
I <sub>0</sub> (μIU/mL)	14.5 (13.0, 17.0)	20.0 (17.0, 24.0)	4.033 × 10 <sup>-3</sup>
I <sub>30</sub> (μIU/mL)	65.0 (52.0, 81.0)	70.5 (52.0, 87.0)	0.8265
I <sub>60</sub> (μIU/mL)	63.0 (50.0, 72.0)	62.5 (49.0, 90.0)	0.9456
I <sub>120</sub> (μIU/mL)	36.5 (27.0, 50.0)	58.5 (47.0, 94.0)	1.17 × 10 <sup>-4</sup>
<b>Other Labs</b>			
Total Cholesterol (mg/dL)	164 (159, 171)	164 (155, 170)	0.6057
HDL Cholesterol (mg/dL)	42 (41, 44)	43 (41, 47)	0.3700
HDL Cholesterol (mg/dL) (M, <i>n</i> = 33, 19)	39 (35, 44)	39 (34, 44)	0.9317
HDL Cholesterol (mg/dL) (F, <i>n</i> = 87, 37)	43 (42, 47)	46 (43, 51)	0.2019
LDL Cholesterol (mg/dL)	94 (92, 104)	101 (94, 105)	0.7447
Triglycerides (mg/dL)	103 (93, 113)	87.5 (78, 114)	0.4017
ALT (IU/L)	18.5 (16, 21)	20 (17, 25)	0.3676
AST (IU/L)	19 (18, 19)	18 (18, 21)	0.8286
TSH (μIU/L)	2.3 (2.0, 2.7)	2.8 (2.3, 3.7)	0.4416
Free T4 (ng/dL)	1.14 (1.1, 1.17)	1.16 (1.13, 1.22)	0.2113
Total T4 (μg/dL)	8.2 (8.0, 8.8)	8.7 (8.3, 9.3)	0.3143

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; CI, confidence interval; G, glucose; HbA1c, hemoglobin A1c; HDL, high-density lipoprotein; HOMA-IR, homeostatic model assessment of insulin resistance; I, insulin; IGM, impaired glucose metabolism; LDL, low-density lipoprotein; NGM, normal glucose metabolism; OGTT, oral glucose tolerance test; TSH, thyroid stimulating hormone; T4, thyroxine.

exact tests were used to compare categorical variables depending on the frequencies of the categories. Two-tailed Student's *t*-tests were used to compare means and 95% confidence intervals of bootstrapped coefficient estimates and scoring metrics.

## 3 Results

### 3.1 Description of data

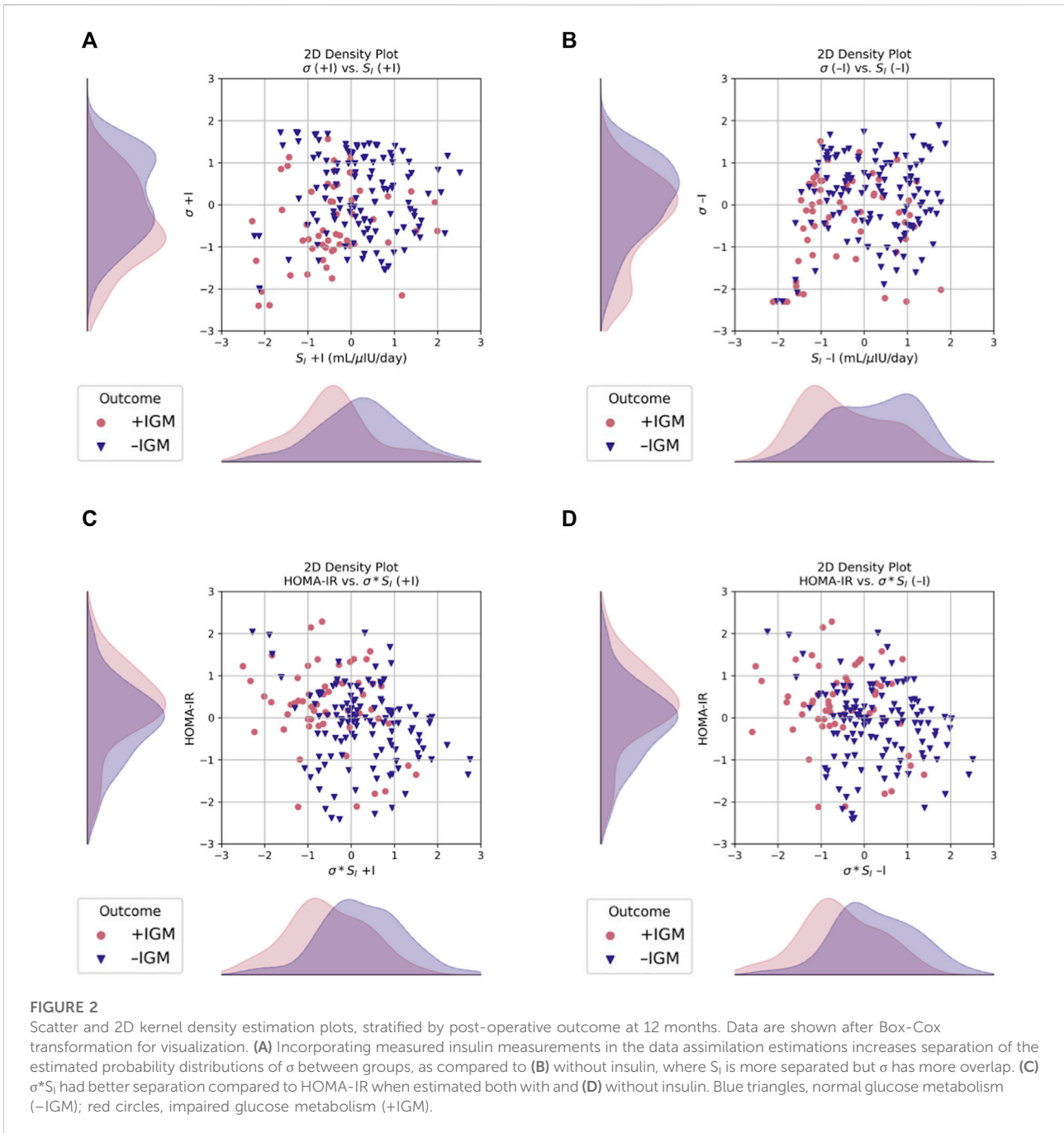
Out of 396 adolescents who underwent bariatric surgery, 248 had pre-operative OGTTs. Of 202 patients without any missing values in their pre-operative OGTT glucose and insulin measurements, 176 had follow-up within the appropriate time window and were further analyzed. All potential classification methods for impaired glucose metabolism (IGM) demonstrated class imbalance, with ~35% (*n* = 56) of patients with meeting post-operative criteria for IGM as outlined above. Pre-operative characteristics by group are shown in Tables 2, 3. Using the same criteria to classify IGM

TABLE 4 Data assimilation-derived mechanistic model parameter estimate comparisons by binary post-operative outcome at 12 months. The parameters  $S_I$ ,  $\sigma$ , and  $\sigma^*S_I$  refer to the means of the posterior distributions for each patient. Overlined parameters were estimated without insulin values. For comparisons, medians and 95% confidence intervals of the median are shown. All comparisons had significant *p*-values for Mann-Whitney-U test after Bonferroni correction (*p* < 0.008).

Parameter Name	– IGM <i>n</i> = 120	+ IGM <i>n</i> = 56	<i>p</i> -value
$\overline{S_I}$	0.403 (0.249, 0.571)	0.086 (0.054, 0.207)	1.70 × 10 <sup>-5</sup>
$S_I$	0.356 (0.284, 0.420)	0.155 (0.129, 0.178)	< 0.0001
$\overline{\sigma^*S_I}$	0.947 (0.651, 1.539)	0.207 (0.157, 0.379)	< 0.0001
$\sigma^*S_I$	0.908 (0.763, 1.510)	0.255 (0.146, 0.547)	< 0.0001
$\overline{\sigma}$	4.715 (4.414, 4.976)	4.083 (3.582, 4.399)	0.0011
$\sigma$	3.424 (2.627, 4.630)	1.596 (1.195, 3.169)	0.0003

Abbreviations: IGM, impaired glucose metabolism;  $\sigma$ , maximal insulin secretion capacity;  $S_I$ , insulin sensitivity.

post-operatively, 101 of the 176 (57.4%) met criteria for IGM pre-operatively. At baseline, patients who had higher probabilities of having IGM post-operatively were more insulin resistant as



measured by HOMA-IR, had higher  $G_{30}$ ,  $G_{60}$ , and  $G_{120}$  values, and had higher  $I_{120}$  values.

A total of 202 patients met inclusion criteria, but 26 of them did not have follow-up within 6–18 months of surgery. One patient was lost to follow-up due to death within three months of surgery. The proportions of patients lost to follow-up were statistically different between those with pre-operative IGM (Fisher’s exact test  $p = 0.035$ ): nine met two or more pre-operative IGM criteria

(three with elevated  $G_0$ , one with elevated  $G_{120}$ , six with multiple elevated HbA1c values, one on anti-diabetic medication).

### 3.2 Data assimilation results

Parameter estimates from data assimilation are summarized in Table 4 and the probability distribution of

TABLE 5 AUROC comparisons between models trained on subsets of features with and without data assimilation. Models are presented in alphabetical order with the better performing model on the left (Model A). Clinical Vars refers to all clinical features in Tables 2, 3 except pre-operative IGM status and HOMA-IR. Clinical Vars<sup>-ins</sup> refers to the same set of features in Clinical Vars after removing insulin measurements. Overlined parameters were estimated without insulin values. Significant *p*-values for two-tailed Student's *t*-test after Bonferroni correction are shown in bold (*p* < 0.004).

Model A	A AUROC Mean (95% CI)	Model B	B AUROC Mean (95% CI)	<i>p</i> -value
1 Clinical Vars	<b>0.7655 (0.7622, 0.7689)</b>	Clinical Vars <sup>-ins</sup> + $\overline{\sigma^*S_I}$	<b>0.7511 (0.7475, 0.7547)</b>	< <b>0.0001</b>
2 Clinical Vars + HOMA-IR	0.7659 (0.7625, 0.7692)	Clinical Vars	0.7655 (0.7622, 0.7689)	0.8942
3 Clinical Vars <sup>-ins</sup> + $\overline{\sigma^*S_I}$	0.7511 (0.7475, 0.7547)	Clinical Vars <sup>-ins</sup>	0.7491 (0.7455, 0.7527)	0.4351
4 Clinical Vars + $\sigma + S_I + \sigma^*S_I$	0.7678 (0.7644, 0.7713)	Clinical Vars	0.7655 (0.7622, 0.7689)	0.3573
5 Clinical Vars + $\sigma + S_I + \sigma^*S_I$	0.7678 (0.7644, 0.7713)	Clinical Vars + HOMA-IR	0.7659 (0.7625, 0.7692)	0.4282
6 Clinical Vars + $\sigma^*S_I$	0.7700 (0.7665, 0.7734)	Clinical Vars	0.7655 (0.7622, 0.7689)	0.0728
7 Clinical Vars + $\sigma^*S_I$	0.7700 (0.7665, 0.7734)	Clinical Vars + HOMA-IR	0.7659 (0.7625, 0.7692)	0.0952
8 <b>Glucose + Insulin + HbA1c + <math>\sigma^*S_I</math></b>	<b>0.7627 (0.7594, 0.7659)</b>	<b>Glucose + Insulin + HbA1c + <math>\sigma + S_I + \sigma^*S_I</math></b>	<b>0.7451 (0.7416, 0.7486)</b>	< <b>0.0001</b>
9 <b>Glucose + Insulin + <math>\sigma + S_I + \sigma^*S_I</math></b>	<b>0.7463 (0.743, 0.7496)</b>	<b>Glucose + Insulin</b>	<b>0.7337 (0.7303, 0.7371)</b>	< <b>0.0001</b>
10 $\overline{\sigma^*S_I}$	0.7380 (0.7346, 0.7415)	Glucose + Insulin	0.7337 (0.7303, 0.7371)	0.0790
11 $\overline{\sigma^*S_I}$	0.7380 (0.7346, 0.7415)	Glucose + Insulin + HOMA-IR	0.7317 (0.7283, 0.7350)	0.0089

Abbreviations: CI, confidence interval; HbA1c, hemoglobin A1c; HOMA-IR, homeostatic model assessment of insulin resistance;  $\sigma$ , maximal insulin secretion capacity;  $S_I$ , insulin sensitivity.

TABLE 6 Average precision comparisons between models trained on subsets of features with and without data assimilation. Models are presented in the same order as in Table 5. Clinical Vars refers to all clinical features in Tables 2, 3 except pre-operative IGM status and HOMA-IR. Clinical Vars<sup>-ins</sup> refers to the same set of features in Clinical Vars after removing insulin measurements. Overlined parameters were estimated without insulin values. Significant *p*-values for two-tailed Student's *t*-test after Bonferroni correction are shown in bold (*p* < 0.004).

Model A	A AP Mean (95% CI)	Model B	B AP Mean (95% CI)	<i>p</i> -value
1 Clinical Vars	0.6200 (0.6148, 0.6252)	Clinical Vars <sup>-ins</sup> + $\overline{\sigma^*S_I}$	0.613 (0.6074, 0.6185)	0.0700
2 Clinical Vars + HOMA-IR	0.6209 (0.6155, 0.6262)	Clinical Vars	0.6200 (0.6148, 0.6252)	0.8200
3 Clinical Vars <sup>-ins</sup> + $\overline{\sigma^*S_I}$	0.613 (0.6074, 0.6185)	Clinical Vars <sup>-ins</sup>	0.6075 (0.6021, 0.6129)	0.1665
4 Clinical Vars + $\sigma + S_I + \sigma^*S_I$	0.6253 (0.6197, 0.6308)	Clinical Vars	0.6200 (0.6148, 0.6252)	0.8835
5 Clinical Vars + $\sigma + S_I + \sigma^*S_I$	0.6253 (0.6197, 0.6308)	Clinical Vars + HOMA-IR	0.6209 (0.6155, 0.6262)	0.2656
6 Clinical Vars + $\sigma^*S_I$	0.6258 (0.6206, 0.6311)	Clinical Vars	0.6200 (0.6148, 0.6252)	0.1244
7 Clinical Vars + $\sigma^*S_I$	0.6258 (0.6206, 0.6311)	Clinical Vars + HOMA-IR	0.6209 (0.6155, 0.6262)	0.1966
8 <b>Glucose + Insulin + HbA1c + <math>\sigma^*S_I</math></b>	<b>0.6156 (0.6105, 0.6207)</b>	<b>Glucose + Insulin + HbA1c + <math>\sigma + S_I + \sigma^*S_I</math></b>	<b>0.5936 (0.5882, 0.599)</b>	< <b>0.0001</b>
9 <b>Glucose + Insulin + <math>\sigma + S_I + \sigma^*S_I</math></b>	<b>0.5841 (0.5791, 0.5892)</b>	<b>Glucose + Insulin</b>	<b>0.5695 (0.5644, 0.5745)</b>	<b>0.0001</b>
10 $\overline{\sigma^*S_I}$	<b>0.5990 (0.5939, 0.6041)</b>	<b>Glucose + Insulin</b>	<b>0.5695 (0.5644, 0.5745)</b>	< <b>0.0001</b>
11 $\overline{\sigma^*S_I}$	<b>0.5990 (0.5939, 0.6041)</b>	<b>Glucose + Insulin + HOMA-IR</b>	<b>0.5662 (0.5611, 0.5713)</b>	< <b>0.0001</b>

Abbreviations: AP, average precision; CI, confidence interval; HbA1c, hemoglobin A1c; HOMA-IR, homeostatic model assessment of insulin resistance;  $\sigma$ , maximal insulin secretion capacity;  $S_I$ , insulin sensitivity.

their means across the cohort as estimated by kernel density estimation are shown in Figure 2. The post-operative + IGM group had statistically significantly lower baseline insulin sensitivity ( $S_I$ ) and maximal insulin secretion capacity ( $\sigma$ ) values compared to the post-operative - IGM group, irrespective of inclusion of insulin in the data assimilation.

The marginal posterior densities of the parameters were estimated for each individual patient using MCMC. The marginal posterior density of  $S_I$  was sharply peaked away from its bounds on manual inspection for a random subset of patients, giving us confidence in its mean estimator. The density for  $\sigma^*S_I$  was also sharply peaked away from its bounds (not shown).

### 3.3 Prediction model results

For the most complex model, data assimilation estimates were negatively correlated with the probability of having post-operative impaired glucose metabolism (IGM) at 12 months. Across models that contained them, HbA1c and  $I_{120}$  were positively correlated with the probability of having post-operative IGM at 12 months. However, the model coefficients were overall not significantly different from zero for all models (not shown). When ranked by magnitude, the largest coefficients, when included, were coefficients for HbA1c,  $I_{120}$ , and  $G_{60}$ .

Selected AUROC and average precision score comparisons are shown in [Tables 5, 6](#), respectively.

The best performing model used all available clinical variables ( $n$  features = 43) and  $\sigma^*S_I$  with an AUROC of 0.77 (95% CI 0.7665, 0.7734) and average precision of 0.6258 (95% CI 0.6206, 0.6311). Our most comprehensive models using all clinical data had similar performances regardless of whether data assimilation estimates or HOMA-IR were included ([Tables 5, 6](#), rows 1–7).

## 4 Discussion

### 4.1 Data assimilation estimates can add clinical information that improves prediction

Our best-performing model used the aforementioned clinical variables combined with the product of maximal insulin secretion capacity and insulin sensitivity,  $\sigma^*S_I$ , with an AUROC of 0.77 (95% CI 0.7665, 0.7734) and average precision of 0.6258 (95% CI 0.6206, 0.6311). This model was nominally better than one using clinical variables alone with an AUROC of 0.7655 (95% CI 0.7622, 0.7689) and average precision of 0.6200 (95% CI 0.6148, 0.6252), but the differences were not significant at  $p = 0.0728$  and 0.1244, respectively ([Tables 5, 6](#), row 6).

### 4.2 Data assimilation estimates can infer missing information encoded in insulin measurements

The comparability of our most comprehensive models suggests that the information added by data assimilation is captured in an extensive, but not exhaustive, clinical dataset. Embedded in the electronic health record (EHR) data was a powerful experiment where the effects of bariatric surgery could be thoroughly investigated. Models incorporating insulin, in general, outperformed models that did not include it. However, when insulin measurements are missing, data assimilation can add physiologic information

that approaches the predictive ability of the full clinical dataset. For example, when using all other clinical features *except* insulin, the model using the product  $\sigma^*S_I$  had nominally improved performance to the model not including any data assimilation estimates, although the  $p$ -value was not significant (AUROC 0.7511 [0.7475, 0.7547] vs. 0.7491 [0.7455, 0.7527], respectively,  $p = 0.4351$ ; AP 0.6130 [0.6074, 0.6185] vs. 0.6075 [0.6021, 0.6129],  $p = 0.1665$ ) ([Tables 5, 6](#), row 3). While the performance of this same model using  $\overline{\sigma^*S_I}$  performed worse than the model using the full clinical dataset including insulin (AUROC 0.7511 [0.7475, 0.7547] vs. 0.7655 [0.7622, 0.7689],  $p < 0.0001$ ; AP 0.6130 [0.6074, 0.6185] vs. 0.6200 [0.6148, 0.6252],  $p = 0.0700$ ) ([Tables 5, 6](#), row 1), its AUROC was non-inferior at 98.5% with  $p = 0.2326$ .

Our models further suggest that  $\sigma^*S_I$ , even when estimated sans insulin, can represent the information within an OGTT using glucose and insulin. When compared, our model trained on only  $\overline{\sigma^*S_I}$  had similar or better performance as compared to models trained on the glucose and insulin measurements from an OGTT (AUROC 0.7380 [0.7346, 0.7415] vs. 0.7337 [0.7303, 0.7371],  $p = 0.079$ ; AP 0.5990 [0.5939, 0.6041] vs. 0.5695 [0.5644, 0.5745],  $p < 0.0001$ ) ([Tables 5, 6](#), row 10). The predictive performance of  $\overline{\sigma^*S_I}$  is demonstrated again in comparison with the model incorporating HOMA-IR, which requires a fasting insulin measurement (AUROC 0.7380 [0.7346, 0.7415] vs. 0.7317 [0.7283, 0.7350],  $p = 0.0089$ ; 0.5990 [0.5939, 0.6041] vs. 0.5662 [0.5611, 0.5713],  $p < 0.0001$ ) ([Tables 5, 6](#), row 11).

### 4.3 The mechanistic models were validated using clinical data

The mechanistic models were able to be well estimated, achieving a stable solution with a unique minimum, using our clinical dataset. Furthermore, the mechanistic model output did not contradict nor add unvalidated information, in that the parameters estimated corresponded to variables associated with glucose and insulin metabolism, and not to other clinical variables (e.g., demographics, thyroid function).

### 4.4 Uncertainty of data assimilation estimates results in reduced performance in logistic regression

When included in the model, HbA1c and insulin measures (particularly  $I_{120}$ ) were frequently ranked as the most important predictors by magnitude. The improved prediction ability using A1c or insulin was not replicated by substituting them with data assimilation estimates not using insulin (not shown). While not a complete substitute

for the information contained in HbA1c or insulin, when maximal insulin secretion capacity ( $\sigma$ ), insulin sensitivity ( $S_I$ ), or  $\sigma^*S_I$  are estimated using insulin measurements, they still improve performance when added to models containing OGTT insulin (AUROC 0.7463 [0.743, 0.7496] vs. 0.7337 [0.7303, 0.7371],  $p < 0.0001$ ; AP 0.5841 [0.5791, 0.5892] vs. 0.5695 [0.5644, 0.5745],  $p = 0.0001$ ) (Tables 5, 6, row 9).

Although in most models the feature coefficients were not statistically different from zero (not shown), the standard deviation of maximal insulin secretion capacity ( $\sigma$ ) is given more importance than the actual estimated parameters themselves and HbA1c in the models using all three data assimilation estimates. Furthermore, in the presence of insulin, regularization consistently shrinks the coefficient for insulin sensitivity ( $S_I$ ), whereas the coefficient for maximal insulin secretion capacity ( $\sigma$ ) increases. The variance in  $\sigma$ 's estimation and its overlapping distributions between groups compared to that of  $S_I$ , even when estimated using insulin measurements, likely contributes to the loss of predictive power (Figure 2). Using  $\sigma^*S_I$  in lieu of  $\sigma$  and  $S_I$  separately improves model performance (AUROC 0.7627 [0.7594, 0.7659] vs. 0.7451 [0.7416, 0.7486],  $p < 0.0001$ ; AP 0.6156 [0.6105, 0.6207] vs. 0.5936 [0.5882, 0.599],  $p < 0.0001$ ) (Tables 5, 6, row 8).

#### 4.5 Disease subtypes can be described by parameter estimates

Differences in pre-operative insulin sensitivity ( $S_I$ ) and maximal insulin secretion capacity ( $\sigma$ ) estimates were seen between outcome groups when estimated both with and without measured insulin ( $p < 0.05$ , Figure 2 and Table 4). Regardless of a patient's pre-operative glycemic status,  $S_I$  better distinguished those patients who would have post-operative impaired glucose metabolism (IGM) compared to either  $\sigma$  or  $\sigma^*S_I$  (Figure 2).

Those with post-operative IGM tended to have lower baseline  $S_I$  and  $\sigma$  values compared to those without it, demonstrating that both may contribute to a patient's disease, although not necessarily equally. Additionally, those with post-operative IGM had higher pre-operative HbA1c,  $G_{30}$ ,  $G_{60}$ , and  $G_{120}$  and  $I_{120}$  values ( $p < 0.05$ , Table 3). This may reflect defects in second phase insulin secretion, which is associated with decreases in insulin secretion capacity (Ha and Sherman, 2020). Notably,  $\sigma^*S_I$  and  $S_I$  values were not correlated with baseline BMI values, demonstrating a seeming disconnect between whole-body adiposity and insulin sensitivity (not shown). Other physiologic parameters that were not estimated using data assimilation in this study, such as insulin secretion rate (ISR) and hepatic insulin sensitivity ( $\text{hepa}S_I$ ), may reveal other disease phenotypes and patient characterizations that could be explored in future work.

#### 4.6 Fixed mechanistic model parameters may differ between adolescents and adults

Not all parameters in the mechanistic models can be estimated simultaneously. All models have potential limitations in generalizability beyond the populations studied during initial development. Use of fixed parameters with values derived from clinical studies in adults may not optimally estimate parameters during pubertal states, where hormonal crosstalk greatly influences energy homeostasis. For example, adolescence is marked by a drastic decrease in insulin sensitivity independent of adiposity; euglycemia is achieved by a compensatory and proportional increase in insulin secretion (Hannon et al., 2006). Insulin secretion rates will therefore be elevated in this age group compared to their adult counterparts, and estimated parameter bounds may differ considerably.

Other parameters which may be relatively constant in adulthood might be more dynamic in adolescents. Obesity and age impact  $\beta$  cell mass and proliferation, which were set as fixed components of these mechanistic models (Saisho et al., 2013; Michaliszyn et al., 2014). The fasting value of  $\gamma$ , representing the  $K^+$ -ATP channel density on the  $\beta$  cells, was set to its example value of  $-0.076$  (Sherman, 2022). However,  $\gamma$  plays an important role in regulating  $\beta$  cell physiology and glucose-mediated insulin secretion. Not properly tuning the fasting value of  $\gamma$  may adversely affect the estimates of the parameters regulating  $\beta$  cell physiology such as maximal insulin secretion capacity ( $\sigma$ ) and insulin priming rate ( $r_2^0$ ). These effects likely reduced the prediction capabilities of our estimates.

Finally, the assumption that estimates based on the 120-min-OGTT approximate parameters as they would be estimated by a hyperinsulinemic euglycemic clamp may be violated in adolescents. Recently, validation of the oral minimal model in adolescents showed that the 120-min-OGTT underestimates insulin sensitivity compared to longer OGTTs, which is not the case for adults (Bartlette et al., 2021). Complicating this suboptimal approximation is the erratic behavior of OGTT measurements, even in euglycemic patients. Glucose and insulin assays can be imprecise, particularly in periods where glucose and insulin are changing rapidly (i.e., after a meal), and removing spurious OGTT measurements can improve performance (Abohtyra et al., 2022).

#### 4.7 Binary classification of outcomes obscures a heterogenous population

Because of the relatively rare frequency of overt diabetes pre-operatively and varying sensitivity of common measurements in adolescents, the post-operative glucose metabolism outcome was coded as a binary variable: normal glucose metabolism (NGM) or

impaired glucose metabolism (IGM). However, there are likely multiple sub-phenotypes present in these groups, which would negatively impact prediction models.

Though the diagnostic threshold for T2DM is HbA1c  $\geq$  6.5%, we used HbA1c  $\geq$  5.7%, the threshold for preDM, to indicate any IGM to improve sensitivity for detecting T2DM in adolescents (Nowicka et al., 2011). The exact interpretation of HbA1c must be considered along with an individual's hemoglobin concentration and structure (Radin, 2014). In particular, studies have shown that traditional HbA1c thresholds for T2DM ( $\geq$  6.5%) are imperfect for diagnosis in adults and may significantly underestimate the prevalence of T2DM in adolescents (Nowicka et al., 2011; American Diabetes Association Professional Practice Committee Draznin et al., 2022). Furthermore, given the frequency of anemia in bariatric surgery patients, HbA1c may misrepresent average blood glucose levels and thus, normal values are insufficient to exclude IGM phenotypes. As such, HbA1c was used as a feature in a logistic regression rather than as the outcome in a linear regression.

All patients were included for analysis regardless of pre-operative IGM status. Notably, neither hyperinsulinism nor insulin resistance were used to code IGM. As insulin resistance is tightly coupled with visceral adiposity, the likelihood that all patients had insulin resistance as the only manifestation of their IGM phenotype is high; this is supported by the lower  $S_1$  values and elevated HOMA-IR calculations of patients at baseline (Kahn and Flier, 2000; Stern et al., 2005). Nonetheless, the decision to focus on glucose rather than insulin perturbations in outcome labeling likely impaired the ability of our models to identify those with IGM-like phenotypes post-operatively, particularly as the outcome is not granular enough to distinguish this heterogenous population.

## 4.8 Limitations

### 4.8.1 Small sample size

Our methods were hindered by a small, imbalanced, and homogenous sample from a single institution. In addition to increasing sample size by using incomplete OGTT data, future work could address class imbalance by under sampling the majority class or removing redundant information. Our sample size relative to features was exacerbated by patient attrition, likely not at random.

### 4.8.2 Bias in clinical measures

Our cohort is more diverse with respect to ethnicity and race compared to previously reported studies, which may impact the ability of a model to predict outcomes when using estimators such as HbA1c. Race differences in outcomes have been seen in other studies and have been attributed to poor calibration of models or measurements across

heterogenous populations (Wallace et al., 2004; Olson et al., 2010; Tharakan et al., 2017). However, race differences between Black and white adolescents have also been noted in hyperinsulinemic-euglycemic clamp studies, so the true difference in T2DM development is still unclear (Michaliszyn et al., 2017).

### 4.8.3 Complexity and validation of electronic health records

Our study did not occur in the context of a clinical trial and is subject to the constraints of EHR data. The most effective use of EHR data applies knowledge of how data are inputted into the system while also understanding the underlying medical decision making process. Improper automatic encoding of features or outcomes could have negatively impacted our models' predictive abilities. In the case of our analyses, features such as comorbidities and drug information are less reliable than laboratory values, and awareness of missing data is not guaranteed—that is, the EHR datasets are not complete (Weiskopf et al., 2013). For example, formal diagnoses of obesity-related liver diseases (i.e., NAFLD and NASH) use a liver biopsy to confirm pathology, which can be done while a patient is undergoing bariatric surgery. However, the diagnosis codes for *suspected* liver disease may not be reflected in the EHR, and if they are inputted, they are done so irregularly. This was the case in our patient sample, where many of the comorbidities were mentioned within the patient assessment and/or radiology impressions of notes, but not listed as a visit-associated diagnosis code frequently enough to be marked as present by our criteria. Relatedly, medication data are sparse and often inaccurate, with unreliable start and stop dates. Diet information, which is undoubtedly important in this context, is not typically represented at all in structured datasets. Information about mental health and psychiatric comorbidities is intentionally difficult to access for secondary use, and social stigmas surrounding mental health reduce confidence that, absent documentation, no comorbidities are present. Similarly, social determinants of health such as food insecurity, exposure to discrimination, and exposure to adverse childhood events are not captured in this dataset. Compounding these limitations are existing health disparities in access bariatric surgery, leading to selection bias (Tsui et al., 2021).

Manual chart review is the gold standard for extracting clinical information, but this is time intensive and subject to human error. It also cannot account for truly missing information. As this relates to our methods, to avoid misclassifying patients with post-operative IGM, we erred on the side of underestimating the proportion of patients with specific comorbidities by requiring at least 25% of encounters to contain the relevant diagnosis codes and structured documentation of medications. This may have classified patients as not having IGM post-operatively when in fact they did have some impaired metabolism.

#### 4.8.4 Limitations related to the use of mechanistic models

The use of mechanistic models and ordinary differential equations inherently limits the number of parameters that can be estimated concurrently. Combined with the limitations related to our retrospective, observational analysis, we are limited in our ability to verify the estimated values with respect to a patient's true physiology. There are several physiologic estimates from data assimilation that could be of use clinically, such as pre-hepatic insulin secretion rate (ISR) or hepatic insulin sensitivity ( $\text{hepa}_{S_i}$ ). Both ISR and  $\text{hepa}_{S_i}$  are difficult to capture clinically, and estimation using OGTTs is complex (Cauter et al., 1992; Kjems et al., 2000). However, our choices to estimate maximal insulin secretion capacity ( $\sigma$ ) and insulin sensitivity ( $S_I$ ) limited our ability to estimate other such important parameters. Generally in our results, more information improved predictions and model accuracy varied in a smooth and logical way. While this does not prove that the ISR estimates incorporated into  $\sigma$  are their true values, it demonstrates that these imperfect features have potential utility in improving patient-level predictions related to surgical outcomes.

Related to the limitations of mechanistic models, our predictive models did not incorporate estimates of hepatic insulin sensitivity ( $\text{hepa}_{S_i}$ ) because  $\text{hepa}_{S_i}$  was necessarily fixed in our equations to better estimate  $\sigma$  and  $S_I$ . However, mismatch between peripheral and hepatic insulin sensitivity may better describe subgroups of adolescents with IGM compared to peripheral insulin resistance alone, particularly in the setting of physiologic pubertal insulin resistance (Hannon et al., 2006). Future work should incorporate this critical component of glucose-insulin metabolism into descriptions of patient phenotypes.

Also absent in these models are the effects of circulating incretins, i.e., glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide or gastric inhibitory polypeptide (GIP). These gut-secreted, insulinotropic hormones are implicated as one potential therapeutic mechanism of action in metabolic surgeries through their actions on insulin secretion and hepatic insulin clearance (Fetner et al., 2005; Hutch and Sandoval, 2017). Supporting this is the general success of novel classes of anti-diabetic drugs leveraging GLP-1 receptor agonists to manage T2DM and obesity in adults and adolescents (Kelly et al., 2020). However, GLP-1 and GIP are not directly represented in the mechanistic models we used, nor is it currently feasible to directly measure their levels in an outpatient clinical setting. Changes in incretins are indirectly reflected in changes in maximal insulin secretion capacity,  $\sigma$ , where increases in GLP-1 lead to increases in  $\sigma$  through the exocytosis model (Ha and Sherman, 2020). If these mechanistic models are employed to predict future physiologic states on a longer time scale, it is critical to include models which incorporate incretin effects to better model post-surgical physiology.

#### 4.9 Future directions

Our model could not capture all features to confidently predict post-surgical glycemic states as a dichotomous outcome despite an extensive collection of laboratory data. We do not believe that these estimates should be broadly disseminated to preclude or exclude patients from receiving indicated care. Rather, with future validation, parameters like  $\sigma$  and  $S_I$  could inform expectations with respect to potential outcomes. Patients intending to have their prediabetic or diabetic states completely reversed should be informed of the possibility that they may not completely resolve with surgery alone. Quantifying that uncertainty may be accomplished using models like the ones described here. By providing more informed consent, we hope that more patients will be able to have meaningful discussions with their care teams to improve long-term surgical outcomes.

Future studies can incorporate using more longitudinal data to see the trends in insulin secretion capacity ( $\sigma$ ) and insulin sensitivity ( $S_I$ ) in the pre-, peri-, and post-operative periods. Applying more granular outcome definitions on a continuous scale may better capture patients who might improve in the severity of their disease, but not sufficiently to resolve impaired glucose metabolism. Alternatively, unsupervised machine learning methods could be applied on a larger cohort of patients to identify different pre-operative phenotypes.

Investigation into additional features that may improve predictive performance can also inform future work. Focusing on the direct effects of surgery itself may provide more insight into patient outcomes, especially when certain surgeries (e.g., sleeve gastrectomy or RYGB) have larger metabolic impacts as compared to relatively metabolically modest restrictive procedures like gastric banding. Using models that directly incorporate the effects of GLP-1 and its secretion in response to glucose ingestion, such as that developed by De Gaetano et al., (De Gaetano et al., 2013) should be included in future work, as should assessment of the change in GLP-1 secretion patterns post-operatively. Prospective studies with larger sample sizes or measurement of more stable insulin byproducts such as C-peptide during OGTTs could improve model performance both by adding a more specific feature and through improvement of data assimilation estimates. Manual chart review can provide information about pre- and post-operative anthropometrics to better quantify adiposity. Alternatively, methods to estimate fat-free body mass using more readily available clinical data could be explored. Inclusion of other candidate biomarkers associated with glucose and insulin metabolism such as incretins, growth factors, inflammatory markers, and carrier proteins could be added. Variables related to behaviors (including diet), mental health, and social determinants of health should also be included in future studies attempting prediction in this same population.

This research is a starting point for further investigation into the use of mechanistic models and data assimilation

applied to clinical problems. In addition to application of similar techniques to clinical problems outside of prediction, research focusing on solving strategies with sparse or irregularly sampled clinical data could provide robust and reliable methods for future studies.

## 5 Conclusion

We demonstrate that data assimilation captures predictive information about glucose metabolism that is not readily apparent from OGTT measurements alone. Further, we validated that our chosen mechanistic model does not add any additional information than it is meant to represent. The clinical variables combined with the product of maximal insulin secretion capacity and insulin sensitivity,  $\sigma^*S_I$ , produced the best-performing model with AUROC = 0.77 and average precision = 0.6258. This model was nominally better than one using clinical variables alone with AUROC = 0.7655, but the difference was not significant at  $p = 0.07$ . In some cases, using the individual components of insulin secretion capacity ( $\sigma$ ) and insulin sensitivity ( $S_I$ ) along with their product reduced prediction model performance.

Looking at whether insulin measurement can be replaced by data assimilation, we found that the model using clinical variables with insulin (AUROC = 0.7655) performed better than the models using clinical variables without insulin but combined with  $\overline{\sigma^*S_I}$  (AUROC = 0.7511,  $p < 0.001$ ). We also found, however, that the difference was small and non-inferior at 98.5%, implying that similar performance can be achieved even without insulin measurements.

If we limit our model inputs to OGTT glucose and insulin values, we found adding data assimilation estimates of insulin secretion capacity and insulin sensitivity ( $\sigma$ ,  $S_I$ , and  $\sigma^*S$ ) significantly increased performance ( $p < 0.001$ ). Models using  $\overline{\sigma^*S_I}$  alone, estimated without insulin, performed marginally better than models using OGTT glucose and insulin with respect to AUROC (0.7380 vs. 0.7337,  $p = 0.08$ ) and had significant improvements in average precision (0.5990 vs. 0.5695,  $p < 0.001$ ).

While data assimilation alone does not significantly improve prediction ability compared to a maximal dataset, the separation of parameter distributions may provide insight into how underlying physiologic processes contribute to a patient's disease. In this adolescent cohort, low insulin sensitivity *and* low maximal insulin secretion capacity distinguish those patients who are less likely to see glycemic benefits from bariatric surgery. While knowing the extent to which defects in glucose-insulin metabolism contribute to disease is not sufficient to confidently predict surgical outcomes, future research can leverage mechanistic models to infer a patient's physiology even when certain data are absent.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author.

## Author contributions

LR, IF, BA, DA, and GH contributed to the conception and design of the study. LR and AO extracted structured and unstructured clinical data from the EHR. IF and JZ provided clinical care for patients in the Center for Adolescent Bariatric Surgery. LR performed manual chart review where necessary. LR cleaned and pre-processed data. BA and DA coded the mechanistic models, and BA performed MCMC fitting. LR and LZ developed the logistic regression models. LR wrote the first draft of the manuscript. BA, LZ, DA, and GH wrote sections of the manuscript. All others contributed to manuscript revision, read, and approved the submitted version.

## Funding

This work was supported by NIH Grant Nos. R01 LM006910 (GH, LZ, AO), R01 LM012734 (GH and DA), and T15 LM007079 (LR and BA).

## Acknowledgments

We acknowledge the excellent support, guidance, and advice from Arthur S. Sherman and Joon Ha to ensure sound use and encoding of their models in this clinical application.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor CDB declared a past co-authorship with authors DA and GH.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



## References

- Abohtyra, R. M., Chan, C. L., Albers, D. J., and Gluckman, B. J. (2022). Inferring insulin secretion rate from sparse patient glucose and insulin measures. *Endocrinol. Incl. Diabetes Mellitus Metabolic Dis.* doi:10.1101/2022.03.10.22272234
- Akalestou, E., Miras, A. D., Rutter, G. A., and le Roux, C. W. (2022). Mechanisms of weight loss after obesity surgery. *Endocr. Rev.* 43, 19–34. doi:10.1210/edrv/bnab022
- Albers, D. J., Blancquart, P.-A., Levine, M. E., Seylbi, E. E., and Stuart, A. (2019). Ensemble kalman methods with constraints. *Inverse Probl.* 35, 095007. doi:10.1088/1361-6420/ab1c09
- Albers, D. J., Levine, M., Gluckman, B., Ginsberg, H., Hripsak, G., and Mamykina, L. (2017). Personalized glucose forecasting for type 2 diabetes using data assimilation. *PLoS Comput. Biol.* 13, e1005232. doi:10.1371/journal.pcbi.1005232
- American Diabetes Association Professional Practice Committee (2022). 2. Classification and diagnosis of diabetes: *Standards of medical Care in diabetes—2022.* *Diabetes Care* 45, S208–S231. doi:10.2337/dc22-S014
- American Diabetes Association Professional Practice Committee (2022). 2. Classification and diagnosis of diabetes: *Standards of medical Care in diabetes—2022.* *Diabetes Care* 45, S17–S38. doi:10.2337/dc22-S002
- Aminian, A., Brethauer, S. A., Andalib, A., Nowacki, A. S., Jimenez, A., Corcelles, R., et al. (2017). Individualized metabolic surgery score: Procedure selection based on diabetes severity. *Ann. Surg.* 266, 650–657. doi:10.1097/SLA.0000000000002407
- Armstrong, S. C., Bolling, C. F., Michalsky, M. P., and Reichard, K. W. (2019). Pediatric metabolic and bariatric surgery: Evidence, barriers, and best practices. *Pediatrics* 144, e20193223. doi:10.1542/peds.2019-3223
- Arslanian, S., El ghormli, L., Young Kim, J., Bacha, F., Chan, C., Ismail, H. M., et al. (2019). The shape of the glucose response curve during an oral glucose tolerance test: Forerunner of heightened glycemic failure rates and accelerated decline in  $\beta$ -cell function in TODAY. *Diabetes Care* 42, 164–172. doi:10.2337/dc18-1122
- Aung, L., Lee, W.-J., Chen, S. C., Ser, K. H., Wu, C. C., Chong, K., et al. (2016). Bariatric surgery for patients with early-onset vs late-onset type 2 diabetes. *JAMA Surg.* 151, 798–805. doi:10.1001/jamasurg.2016.1130
- Bal, B. S., Finelli, F. C., Shope, T. R., and Koch, T. R. (2012). Nutritional deficiencies after bariatric surgery. *Nat. Rev. Endocrinol.* 8, 544–556. doi:10.1038/nrendo.2012.48
- Bartlette, K., Carreau, A.-M., Xie, D., Garcia-Reyes, Y., Rahat, H., Pyle, L., et al. (2021). Oral minimal model-based estimates of insulin sensitivity in obese youth depend on oral glucose tolerance test protocol duration. *Metabol. Open* 9, 100078. doi:10.1016/j.metop.2021.100078
- Bartolomé, A., Suda, N., Yu, J., Zhu, C., Son, J., Ding, H., et al. (2022). Notch-mediated Ephrin signaling disrupts islet architecture and  $\beta$  cell function. *JCI Insight* 7, e157694. doi:10.1172/jci.insight.157694
- Beamish, A. J., and Reinher, T. (2017). Should bariatric surgery be performed in adolescents? *Eur. J. Endocrinol.* 176, D1–D15. D1–D15. doi:10.1530/EJE-16-0906
- Bergman, R. N., Prager, R., Volund, A., and Olefsky, J. M. (1987). Equivalence of the insulin sensitivity index in man derived by the minimal model method and the euglycemic glucose clamp. *J. Clin. Invest.* 79, 790–800. doi:10.1172/JCI112886
- Bolling, C. F., Armstrong, S. C., Reichard, K. W., and Michalsky, M. P. (2019). Metabolic and bariatric surgery for pediatric patients with severe obesity. *Pediatrics* 144, e20193224. doi:10.1542/peds.2019-3224
- Box, G. E. P., and Cox, D. R. (1964). An analysis of transformations. *J. R. Stat. Soc. Ser. B* 26, 211–243. doi:10.1111/j.2517-6161.1964.tb00553.x
- Brown, R. J., and Yanovski, J. A. (2014). Estimation of insulin sensitivity in children: Methods, measures and controversies. *Pediatr. Diabetes* 15, 151–161. doi:10.1111/peidi.12146
- Bryant, M., Ashton, L., Brown, J., Jebb, S., Wright, J., Roberts, K., et al. (2014). Systematic review to identify and appraise outcome measures used to evaluate childhood obesity treatment interventions (CoOR): Evidence of purpose, application, validity, reliability and sensitivity. *Health Technol. Assess.* 18, 1–380. doi:10.3310/hta18510
- Buck, S. F. (1960). A method of estimation of missing values in multivariate data suitable for use with an electronic computer. *J. R. Stat. Soc. Ser. B* 22, 302–306. doi:10.1111/j.2517-6161.1960.tb00375.x
- Buse, J. B., Caprio, S., Cefalu, W. T., Ceriello, A., Del Prato, S., Inzucchi, S. E., et al. (2009). How do we define cure of diabetes? *Diabetes Care* 32, 2133–2135. doi:10.2337/dc09-9036
- Buse, J. B., Kaufman, F. R., Linder, B., Hirst, K., El Ghormli, L., Willi, S., et al. (2013). Diabetes screening with hemoglobin A1c versus fasting plasma glucose in a multiethnic middle-school cohort. *Diabetes Care* 36, 429–435. doi:10.2337/dc12-0295
- Buuren, S. van, and Groothuis-Oudshoorn, K. (2011). Mice : Multivariate imputation by chained equations in R. *J. Stat. Softw.* 45. doi:10.18637/jss.v045.i03
- Cao, Y., Näslund, I., Näslund, E., Ottosson, J., Montgomery, S., and Stenberg, E. (2020). Using convolutional neural network to predict remission of diabetes after gastric bypass surgery: A machine learning study from the scandinavian obesity surgery register. *Endocrinol. Incl. Diabetes Mellitus Metabolic Dis.* doi:10.1101/2020.11.03.20224956
- Cauter, E. V., Mestrez, F., Sturis, J., and Polonsky, K. S. (1992). Estimation of insulin secretion rates from C-peptide levels: Comparison of individual and standard kinetic parameters for C-peptide clearance. *Diabetes* 41, 368–377. doi:10.2337/diab.41.3.368
- Chen, Y., Wang, S., and Sherman, A. (2008). Identifying the targets of the amplifying pathway for insulin secretion in pancreatic  $\beta$ -cells by kinetic modeling of granule exocytosis. *Biophys. J.* 95, 2226–2241. doi:10.1529/biophysj.107.124990
- Chung, S. T., Katz, L. E. L., Stettler-Davis, N., Shults, J., Sherman, A., Ha, J., et al. (2022). The relationship between lipoproteins and insulin sensitivity in youth with obesity and abnormal glucose tolerance. *J. Clin. Endocrinol. Metabolism* 107, 1541–1551. doi:10.1210/clinem/dgac113
- Cobelli, C., Dalla Man, C., Toffolo, G., Basu, R., Vella, A., and Rizza, R. (2014). The oral minimal model method. *Diabetes* 63, 1203–1213. doi:10.2337/db13-1198
- Courcoulas, A. P., Christian, N. J., Belle, S. H., Berk, P. D., Flum, D. R., Garcia, L., et al. (2013). Weight change and health outcomes at 3 Years after bariatric surgery among individuals with severe obesity. *JAMA* 310, 2416–2425. doi:10.1001/jama.2013.280928
- Courcoulas, A. P., Goodpaster, B. H., Eagleton, J. K., Belle, S. H., Kalarchian, M. A., Lang, W., et al. (2014). Surgical vs medical treatments for type 2 diabetes mellitus: A randomized clinical trial. *JAMA Surg.* 149, 707–715. doi:10.1001/jamasurg.2014.467
- Crossan, K., and Sheer, A. J. (2022). “Surgical options in the treatment of severe obesity,” in StatPearls. *StatPearls publishing: Treasure island, FL.* Available at: <https://www.ncbi.nlm.nih.gov/books/NBK576372/> (accessed Sep 13 2022).
- Dabelea, D., Mayer-Davis, E. J., Saydah, S., Imperatore, G., Linder, B., Divers, J., et al. (2014). Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *JAMA* 311, 1778–1786. doi:10.1001/jama.2014.3201
- Dalla Man, C., Rizza, R. A., and Cobelli, C. (2007). Meal simulation model of the glucose-insulin system. *IEEE Trans. Biomed. Eng.* 54, 1740–1749. doi:10.1109/TBME.2007.893506
- De Gaetano, A., Panunzi, S., Matone, A., Samson, A., Vrbikova, J., Bendlova, B., et al. (2013). Routine OGTT: A robust model including incretin effect for precise identification of insulin sensitivity and secretion in a single individual. *PLoS ONE* 8, e70875. doi:10.1371/journal.pone.0070875
- DeMaria, E. J., Portenier, D., and Wolfe, L. (2007). Obesity surgery mortality risk score: Proposal for a clinically useful score to predict mortality risk in patients undergoing gastric bypass. *Surg. Obes. Relat. Dis.* 3, 134–140. doi:10.1016/j.soard.2007.01.005
- Dewberry, L. C., Niemiec, S. M., Hilton, S. A., Louiselle, A. E., Singh, S., Sakthivel, T. S., et al. (2022). Cerium oxide nanoparticle conjugation to microRNA-146a mechanism of correction for impaired diabetic wound healing. *Nanomedicine*. 40, 102483. doi:10.1016/j.nano.2021.102483
- Dixon, J. B., Chuang, L.-M., Chong, K., Chen, S. C., Lambert, G. W., Straznicki, N. E., et al. (2013). Predicting the glycemic response to gastric bypass surgery in patients with type 2 diabetes. *Diabetes Care* 36, 20–26. doi:10.2337/dc12-0779
- Elbahrawy, A., Bougie, A., Loiselle, S.-E., Demyttenaere, S., Court, O., and Andalib, A. (2018). Medium to long-term outcomes of bariatric surgery in older adults with super obesity. *Surg. Obes. Relat. Dis.* 14, 470–476. doi:10.1016/j.soard.2017.11.008
- Evensen, G. (2009). *Data assimilation: The ensemble kalman filter.* Berlin, Heidelberg: Springer Berlin Heidelberg Springer e-books.
- Fetner, R., McGinty, J., Russell, C., Pi-Sunyer, F. X., and Laferrère, B. (2005). Incretins, diabetes, and bariatric surgery: A review. *Surg. Obes. Relat. Dis.* 1, 589–597. doi:10.1016/j.soard.2005.09.001
- Ge, H., Xu, K., and Ghahramani, Z. (2018). Turing: A language for flexible probabilistic inference. In *Int. Conf. Artif. Intell. Statistics*, 9–11. Spain: AISTATS, 1682–1690.

- Gunczler, P., and Lanes, R. (2006). Relationship between different fasting-based insulin sensitivity indices in obese children and adolescents. *J. Pediatr. Endocrinol. Metab.* 19, 259–265. doi:10.1515/JPEM.2006.19.3.259
- Gutch, M., Kumar, S., Razi, S. M., Gupta, K. K., and Gupta, A. (2015). Assessment of insulin sensitivity/resistance. *Indian J. Endocrinol. Metab.* 19, 160–164. doi:10.4103/2230-8210.146874
- Ha, J., Satin, L. S., and Sherman, A. S. (2016). A mathematical model of the pathogenesis, prevention, and reversal of type 2 diabetes. *Endocrinology* 157, 624–635. doi:10.1210/en.2015-1564
- Ha, J., and Sherman, A. (2020). Type 2 diabetes: One disease, many pathways. *Am. J. Physiol. Endocrinol. Metab.* 319, E410–E426–E426. doi:10.1152/ajpendo.00512.2019
- Hales, C. M., Carroll, M. D., Fryar, C. D., and Ogden, C. L. (2017h). *Prevalence of obesity among adults and youth: United States, 2015–2016*. Hyattsville, MD: National Center for Health Statistics. [www.cdc.gov/nchs/data/databriefs/db288.pdf](http://www.cdc.gov/nchs/data/databriefs/db288.pdf).
- Hannon, T. S., Janosky, J., and Arslanian, S. A. (2006). Longitudinal study of physiologic insulin resistance and metabolic changes of puberty. *Pediatr. Res.* 60, 759–763. doi:10.1203/01.pdr.0000246097.73031.27
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 97–109. doi:10.1093/biomet/57.1.97
- Hatoum, I. J., Greenawald, D. M., Cotsapas, C., Reitman, M. L., Daly, M. J., and Kaplan, L. M. (2011). Heritability of the weight loss response to gastric bypass surgery. *J. Clin. Endocrinol. Metab.* 96, E1630–E1633. doi:10.1210/jc.2011-1130
- Holst, J. J., and Madsbad, S. (2021). What is diabetes remission? *Diabetes Ther.* 12, 641–646. doi:10.1007/s13300-021-01032-y
- Hsia, D. S., Fallon, S. C., and Brandt, M. L. (2012). Adolescent bariatric surgery. *Arch. Pediatr. Adolesc. Med.* 166, 757–766. doi:10.1001/archpediatrics.2012.1011
- Hutch, C. R., and Sandoval, D. (2017). The role of GLP-1 in the metabolic success of bariatric surgery. *Endocrinology* 158, 4139–4151. doi:10.1210/en.2017-00564
- Inge, T. H., Coley, R. Y., Bazzano, L. A., Xanthakos, S. A., McTigue, K., Arterburn, D., et al. (2018). Comparative effectiveness of bariatric procedures among adolescents: The PCORnet bariatric study. *Surg. Obes. Relat. Dis.* 14, 1374–1386. doi:10.1016/j.soard.2018.04.002
- Inge, T. H., Courcoulas, A. P., Jenkins, T. M., Michalsky, M. P., Brandt, M. L., Xanthakos, S. A., et al. (2019). Five-year outcomes of gastric bypass in adolescents as compared with adults. *N. Engl. J. Med.* 380, 2136–2145. doi:10.1056/NEJMoa1813909
- Inge, T. H., Courcoulas, A. P., Jenkins, T. M., Michalsky, M. P., Helmrath, M. A., Brandt, M. L., et al. (2016). Weight loss and health status 3 Years after bariatric surgery in adolescents. *N. Engl. J. Med.* 374, 113–123. doi:10.1056/NEJMoa1506699
- Inge, T. H., Jenkins, T. M., Xanthakos, S. A., Dixon, J. B., Daniels, S. R., Zeller, M. H., et al. (2017). Long-term outcomes of bariatric surgery in adolescents with severe obesity (FABS-5+): A prospective follow-up analysis. *Lancet. Diabetes Endocrinol.* 5, 165–173. doi:10.1016/S2213-8587(16)30315-1
- Inge, T. H., Laffel, L. M., Jenkins, T. M., Marcus, M. D., Leibel, N. I., Brandt, M. L., et al. (2018). Comparison of surgical and medical therapy for type 2 diabetes in severely obese adolescents. *JAMA Pediatr.* 172, 452–460. doi:10.1001/jamapediatrics.2017.5763
- Inge, T. H., Zeller, M. H., Jenkins, T. M., Helmrath, M., Brandt, M. L., Michalsky, M. P., et al. (2014). Perioperative outcomes of adolescents undergoing bariatric surgery: The teen-longitudinal assessment of bariatric surgery (Teen-LABS) study. *JAMA Pediatr.* 168, 47–53. doi:10.1001/jamapediatrics.2013.4296
- Julier, S. J., and Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. *Proc. IEEE* 92, 401–422. doi:10.1109/jproc.2003.823141
- Kahn, B. B., and Flier, J. S. (2000). Obesity and insulin resistance. *J. Clin. Investig.* 106, 473–481. doi:10.1172/JCI10842
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *J. Basic Eng.* 82, 35–45. doi:10.1115/1.3662552
- Kam, H., Tu, Y., Pan, J., Han, J., Zhang, P., Bao, Y., et al. (2020). Comparison of four risk prediction models for diabetes remission after roux-en-Y gastric bypass surgery in obese Chinese patients with type 2 diabetes mellitus. *Obes. Surg.* 30, 2147–2157. doi:10.1007/s11695-019-04371-9
- Karasko, D. (2019). Weight loss in adolescents after bariatric surgery: A systematic review. *J. Pediatr. Health Care* 33, 26–34. doi:10.1016/j.pedhc.2018.05.010
- Keidar, A. (2011). Bariatric surgery for type 2 diabetes reversal: The risks. *Diabetes Care* 34, S361–S266. doi:10.2337/dc11-s254
- Kelly, A. S., Auerbach, P., Barrientos-Perez, M., Gies, L., Hale, P. M., Marcus, C., et al. (2020). A randomized, controlled trial of liraglutide for adolescents with obesity. *N. Engl. J. Med.* 382, 2117–2128. doi:10.1056/NEJMoa1916038
- Khattab, A., and Sperling, M. A. (2019). Obesity in adolescents and youth: The case for and against bariatric surgery. *J. Pediatr.* 207, 18–22. doi:10.1016/j.jpeds.2018.11.058
- Kim, J. Y., Tfayli, H., Bacha, F., Lee, S., Michalyszyn, S. F., Yousuf, S., et al. (2020).  $\beta$ -cell function, incretin response, and insulin sensitivity of glucose and fat metabolism in obese youth: Relationship to OGTT-time-to-glucose-peak. *Pediatr. Diabetes* 21, 18–27. doi:10.1111/pedi.12940
- Kjems, L. L., Christiansen, E., Volund, A., Bergman, R. N., Madsbad, S., and Volund, A. (2000). Validation of methods for measurement of insulin secretion in humans *in vivo*. *Diabetes* 49, 580–588. doi:10.2337/diabetes.49.4.580
- Knowler, W. C., Barrett-Connor, E., Fowler, S. E., Hamman, R. F., Lachin, J. M., Walker, E. A., et al. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N. Engl. J. Med.* 346, 393–403. doi:10.1056/NEJMoa012512
- Laferrère, B., and Pattou, F. (2018). Weight-independent mechanisms of glucose control after roux-en-Y gastric bypass. *Front. Endocrinol.* 9, 530. doi:10.3389/fendo.2018.00530
- Lamoshi, A., Chernoguz, A., Harmon, C. M., and Helmrath, M. (2020). Complications of bariatric surgery in adolescents. *Semin. Pediatr. Surg.* 29, 150888. doi:10.1016/j.sempedsurg.2020.150888
- Law, K. J. H., Stuart, A. M., and Zygalakis, K. C. (2015). *Data assimilation: A mathematical introduction*. arXiv:150607825 [math, stat] Available at: <http://arxiv.org/abs/1506.07825> (accessed Mar 15, 2022).
- Lee, J. M. (2007). Insulin resistance in children and adolescents. *Rev. Endocr. Metab. Disord.* 7, 141–147. doi:10.1007/s11154-006-9019-8
- Lee, J. M., Okumura, M. J., Davis, M. M., Herman, W. H., and Gurney, J. G. (2006). Prevalence and determinants of insulin resistance among U.S. Adolescents: A population-based study. *Diabetes Care* 29, 2427–2432. doi:10.2337/dc06-0709
- Levine, M. E., Hripsak, G., Mamykina, L., Stuart, A., and Albers, D. J. (2017). *Offline and online data assimilation for real-time blood glucose forecasting in type 2 diabetes*. arXiv:170900163 [math, q-bio] Available at: <http://arxiv.org/abs/1709.00163> (accessed Mar 15, 2022).
- Lim, E. L., Hollingsworth, K. G., Aribisala, B. S., Chen, M. J., Mathers, J. C., and Taylor, R. (2011). Reversal of type 2 diabetes: Normalisation of beta cell function in association with decreased pancreas and liver triacylglycerol. *Diabetologia* 54, 2506–2514. doi:10.1007/s00125-011-2204-7
- Little, R. R., Rohlfing, C. L., Tennill, A. L., Madsen, R. W., Polonsky, K. S., Myers, G. L., et al. (2008). Standardization of C-peptide measurements. *Clin. Chem.* 54, 1023–1026. doi:10.1373/clinchem.2007.101287
- Livhits, M., Mercado, C., Yermilov, I., Parikh, J. A., Dutson, E., Mehran, A., et al. (2012). Preoperative predictors of weight loss following bariatric surgery: Systematic review. *Obes. Surg.* 22, 70–89. doi:10.1007/s11695-011-0472-4
- Manley, S. E., Stratton, I. M., Clark, P. M., and Luzio, S. D. (2007). Comparison of 11 human insulin assays: Implications for clinical investigation and research. *Clin. Chem.* 53, 922–932. doi:10.1373/clinchem.2006.077784
- Matthews, D. R., Hosker, J. P., Rudenski, A. S., Naylor, B. A., Treacher, D. F., and Turner, R. C. (1985). Homeostasis model assessment: Insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28, 412–419. doi:10.1007/BF00280883
- Michalyszyn, S. F., Lee, S., Bacha, F., Tfayli, H., Farchoukh, L., Mari, A., et al. (2017). Differences in  $\beta$ -cell function and insulin secretion in Black vs. White obese adolescents: Do incretin hormones play a role? Race,  $\beta$ -cell function, and incretin response. *Pediatr. Diabetes* 18, 143–151. doi:10.1111/pedi.12364
- Michalyszyn, S. F., Mari, A., Lee, S., Bacha, F., Tfayli, H., Farchoukh, L., et al. (2014).  $\beta$ -Cell function, incretin effect, and incretin hormones in obese youth along the span of glucose tolerance from normal to prediabetes to type 2 diabetes. *Diabetes* 63, 3846–3855. doi:10.2337/db13-1951
- Miller, W. G., Thienpont, L. M., Van Uytanghe, K., Clark, P. M., Lindstedt, P., Nilsson, G., et al. (2009). Toward standardization of insulin immunoassays. *Clin. Chem.* 55, 1011–1018. doi:10.1373/clinchem.2008.118380
- Mingrone, G., Panunzi, S., De Gaetano, A., Guidone, C., Iaconelli, A., Nanni, G., et al. (2015). Bariatric-metabolic surgery versus conventional medical treatment in obese patients with type 2 diabetes: 5 year follow-up of an open-label, single-centre, randomised controlled trial. *Lancet* 386, 964–973. doi:10.1016/S0140-6736(15)00075-6
- Montero, P. N., Stefanidis, D., Norton, H. J., Gersin, K., and Kuwada, T. (2011). Reported excess weight loss after bariatric surgery could vary significantly depending on calculation method: A plea for standardization. *Surg. Obes. Relat. Dis.* 7, 531–534. doi:10.1016/j.soard.2010.09.025
- Mulgrave, J. J., Levine, M. E., et al. (2020). *Using data assimilation of mechanistic models to estimate glucose and insulin metabolism*. arXiv:200306541 [physics, stat] Available at: <http://arxiv.org/abs/2003.06541> (accessed Mar 16, 2021).
- Mulla, C. M., Middelbeek, R. J. W., and Patti, M.-E. (2018). Mechanisms of weight loss and improved metabolism following bariatric surgery. *Ann. N. Y. Acad. Sci.* 1411, 53–64. doi:10.1111/nyas.13409

- Muniyappa, R., and Madan, R. (2018). "Assessing insulin sensitivity and resistance in humans," in *Endotext*. Editors K. Feingold, B. Anawalt, and A. Boyce (South Dartmouth, MA. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK278954/>).
- Nathan, D. M., Davidson, M. B., DeFronzo, R. A., Heine, R. J., Henry, R. R., Pratley, R., et al. (2007). Impaired fasting glucose and impaired glucose tolerance: Implications for care. *Diabetes Care* 30, 753–759. doi:10.2337/dc07-9920
- Nowicka, P., Santoro, N., Liu, H., Lartaud, D., Shaw, M. M., Goldberg, R., et al. (2011). Utility of hemoglobin A(1c) for diagnosing prediabetes and diabetes in obese children and adolescents. *Diabetes Care* 34, 1306–1311. doi:10.2337/dc10-1984
- Okser, S., Pahikkala, T., and Aittokallio, T. (2013). Genetic variants and their interactions in disease risk prediction – machine learning and network perspectives. *BioData Min.* 6, 5. doi:10.1186/1756-0381-6-5
- Olbers, T., Beamish, A. J., Gronowitz, E., Flodmark, C. E., Dahlgren, J., Bruze, G., et al. (2017). Laparoscopic roux-en-Y gastric bypass in adolescents with severe obesity (AMOS): A prospective, 5-year, Swedish nationwide study. *Lancet. Diabetes Endocrinol.* 5, 174–183. doi:10.1016/S2213-8587(16)30424-7
- Olson, D. E., Rhee, M. K., Herrick, K., Ziemer, D. C., Twombly, J. G., and Phillips, L. S. (2010). Screening for diabetes and pre-diabetes with proposed A1C-based diagnostic criteria. *Diabetes Care* 33, 2184–2189. doi:10.2337/dc10-0433
- Ortega, E., Morinigo, R., Flores, L., Moize, V., Rios, M., Lacy, A. M., et al. (2012). Predictive factors of excess body weight loss 1 year after laparoscopic bariatric surgery. *Surg. Endosc.* 26, 1744–1750. doi:10.1007/s00464-011-2104-4
- Ou, X., Chen, M., Xu, L., Lin, W., Huang, H., Chen, G., et al. (2022). Changes in bone mineral density after bariatric surgery in patients of different ages or patients with different postoperative periods: A systematic review and meta-analysis. *Eur. J. Med. Res.* 27, 144. doi:10.1186/s40001-022-00774-0
- Pajvani, U. B., and Accili, D. (2015). The new biology of diabetes. *Diabetologia* 58, 2459–2468. doi:10.1007/s00125-015-3722-5
- Panunzi, S., De Gaetano, A., Carnicelli, A., and Mingrone, G. (2015). Predictors of remission of diabetes mellitus in severely obese individuals undergoing bariatric surgery: Do BMI or procedure choice matter? A meta-analysis. *Ann. Surg.* 261, 459–467. doi:10.1097/SLA.0000000000000863
- Pedersen, H. K., Gudmundsdottir, V., Pedersen, M. K., Brorsson, C., Brunak, S., and Gupta, R. (2016). Ranking factors involved in diabetes remission after bariatric surgery using machine-learning integrating clinical and genomic biomarkers. *NPJ Genom. Med.* 1, 16035. doi:10.1038/npjgenmed.2016.35
- Pedregosa, F., Varoquaux, G., and Gramfort, A. (2011). Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Pedroso, F. E., Angriman, F., Endo, A., Dasenbrock, H., Storino, A., Castillo, R., et al. (2018). Weight loss after bariatric surgery in obese adolescents: A systematic review and meta-analysis. *Surg. Obes. Relat. Dis.* 14, 413–422. doi:10.1016/j.soard.2017.10.003
- Perotte, A., and Hripcsak, G. (2013). Temporal properties of diagnosis code time series in aggregate. *IEEE J. Biomed. Health Inf.* 17, 477–483. doi:10.1109/JBHI.2013.2244610
- Prentki, M., and Nolan, C. J. (2006). Islet cell failure in type 2 diabetes. *J. Clin. Invest.* 116, 1802–1812. doi:10.1172/JCI29103
- Purnell, J. Q., Dewey, E. N., Laferrère, B., Selzer, F., Flum, D. R., Mitchell, J. E., et al. (2021). Diabetes remission status during seven-year follow-up of the longitudinal assessment of bariatric surgery study. *J. Clin. Endocrinol. Metab.* 106, 774–788. doi:10.1210/clinem/dgab849
- Rackauckas, C., and Nie, Q. (2017). Differentialequations.jl—a performant and feature-rich ecosystem for solving differential equations in julia. *J. Open Res. Softw.* 5, 15. doi:10.5334/jors.151
- Radin, M. S. (2014). Pitfalls in hemoglobin A1c measurement: When results may be misleading. *J. Gen. Intern. Med.* 29, 388–394. doi:10.1007/s11606-013-2595-x
- Ramos-Levi, A., Sanchez-Pernaute, A., Matia, P., Cabrerizo, L., Barabash, A., Hernandez, C., et al. (2013). Diagnosis of diabetes remission after bariatric surgery may be jeopardized by remission criteria and previous hypoglycemic treatment. *Obes. Surg.* 23, 1520–1526. doi:10.1007/s11695-013-0995-y
- Ramos-Levi, A., Sanchez-Pernaute, A., Matia, P., Cabrerizo, L., Barabash, A., Hernandez, C., et al. (2013). Diagnosis of diabetes remission after bariatric surgery may be jeopardized by remission criteria and previous hypoglycemic treatment. *Obes. Surg.* 23, 1520–1526. doi:10.1007/s11695-013-0995-y
- Richter, L. R., Wan, Q., Wen, D., Zhang, Y., Yu, J., Kang, J. K., et al. (2020). Targeted delivery of notch inhibitor attenuates obesity-induced glucose intolerance and liver fibrosis. *ACS Nano* 14, 6878–6886. doi:10.1021/acsnano.0c01007
- Robert, M., Ferrand-Gaillard, C., Disse, E., Espalieu, P., Simon, C., Laville, M., et al. (2013). Predictive factors of type 2 diabetes remission 1 Year after bariatric surgery: Impact of surgical techniques. *Obes. Surg.* 23, 770–775. doi:10.1007/s11695-013-0868-4
- Rouskas, K., Cauchi, S., Raverdy, V., Yengo, L., Froguel, P., and Pattou, F. (2014). Weight loss independent association of TCF7 L2 gene polymorphism with fasting blood glucose after Roux-en-Y gastric bypass in type 2 diabetic patients. *Surg. Obes. Relat. Dis.* 10, 679–683. doi:10.1016/j.soard.2013.12.016
- Roy, V. (2019). *Convergence diagnostics for Markov chain Monte Carlo*. arXiv:190911827 [stat] Available at: <http://arxiv.org/abs/1909.11827> (accessed Apr 13, 2022).
- Rubino, F., Nathan, D. M., Eckel, R. H., Schauer, P. R., Alberti, K. G. M. M., Zimmet, P. Z., et al. (2016). Metabolic surgery in the treatment algorithm for type 2 diabetes: A joint statement by international diabetes organizations. *Diabetes Care* 39, 861–877. doi:10.2337/dc16-0236
- Saisho, Y., Butler, A. E., Manesso, E., Elashoff, D., Rizza, R. A., and Butler, P. C. (2013).  $\beta$ -Cell mass and turnover in humans: Effects of obesity and aging. *Diabetes Care* 36, 111–117. doi:10.2337/dc12-0421
- Schauer, P. R., Kashyap, S. R., Wolski, K., Brethauer, S. A., Kirwan, J. P., Pothier, C. E., et al. (2012). Bariatric surgery versus intensive medical therapy in obese patients with diabetes. *N. Engl. J. Med.* 366, 1567–1576. doi:10.1056/NEJMoa1200225
- Seeley, R. J., Chambers, A. P., and Sandoval, D. A. (2015). The role of gut adaptation in the potent effects of multiple bariatric surgeries on obesity and diabetes. *Cell. Metab.* 21, 369–378. doi:10.1016/j.cmet.2015.01.001
- Shaibi, G. Q., Davis, J. N., Weigensberg, M. J., and Goran, M. I. (2011). Improving insulin resistance in obese youth: Choose your measures wisely. *Int. J. Pediatr. Obes.* 6, e290–e296. doi:10.3109/17477166.2010.528766
- Shen, S.-C., Wang, W., Tam, K.-W., Chen, H. A., Lin, Y. K., Wang, S. Y., et al. (2019). Validating risk prediction models of diabetes remission after sleeve gastrectomy. *Obes. Surg.* 29, 221–229. doi:10.1007/s11695-018-3510-7
- Sherman, A. (2022). *Mathematical model of diabetes pathways*. doi:10.6084/m9.figshare.10792412.v9
- Sinha, R., Fisch, G., Teague, B., Tamborlane, W. V., Banyas, B., Allen, K., et al. (2002). Prevalence of impaired glucose tolerance among children and adolescents with marked obesity. *N. Engl. J. Med.* 346, 802–810. doi:10.1056/NEJMoa012578
- Skellton, J. A., Cook, S. R., Auinger, P., Klein, J. D., and Barlow, S. E. (2009). Prevalence and trends of severe obesity among US children and adolescents. *Acad. Pediatr.* 9, 322–329. doi:10.1016/j.acap.2009.04.005
- Stefater, M. A., and Inge, T. H. (2017). Bariatric surgery for adolescents with type 2 diabetes: An emerging therapeutic strategy. *Curr. Diab. Rep.* 17, 62. doi:10.1007/s11892-017-0887-y
- Stern, S. E., Williams, K., Ferrannini, E., DeFronzo, R. A., Bogardus, C., and Stern, M. P. (2005). Identification of individuals with insulin resistance using routine clinical measurements. *Diabetes* 54, 333–339. doi:10.2337/diabetes.54.2.333
- Sturis, J., Polonsky, K. S., Mosekilde, E., and Van Cauter, E. (1991). Computer model for mechanisms underlying ultradian oscillations of insulin and glucose. *Am. J. Physiol.* 260, E801–E809. doi:10.1152/ajpendo.1991.260.5.E801
- Tagi, V. M., Giannini, C., and Chiarelli, F. (2019). Insulin resistance in children. *Front. Endocrinol.* 10, 342. doi:10.3389/fendo.2019.00342
- Taylor, R., Al-Mrabeh, A., and Sattar, N. (2019). Understanding the mechanisms of reversal of type 2 diabetes. *Lancet. Diabetes Endocrinol.* 7, 726–736. doi:10.1016/S2213-8587(19)30076-2
- Tharakan, G., Scott, R., Szepletowski, O., Miras, A. D., Blakemore, A. I., Purkayastha, S., et al. (2017). Limitations of the DiaRem score in predicting remission of diabetes following roux-en-Y gastric bypass (RYGB) in an ethnically diverse population from a single institution in the UK. *Obes. Surg.* 27, 782–786. doi:10.1007/s11695-016-2368-9
- TODAY Study GroupBjornstad, P., Drews, K. L., Caprio, S., Gubitosi-Klug, R., and Nathan, D. M. (2021). Long-term complications in youth-onset type 2 diabetes. *N. Engl. J. Med.* 385, 416–426. doi:10.1056/NEJMoa2100165
- Toh, S., Rasmussen-Torvik, L. J., Harmata, E. E., Pardee, R., Saizan, R., Malanga, E., et al. (2017). The national patient-centered clinical research network (PCORnet) bariatric study cohort: Rationale, methods, and baseline characteristics. *JMIR Res. Protoc.* 6, e222. doi:10.2196/resprot.8323
- Tohidi, M., Arbab, P., and Ghasemi, A. (2017). Assay-dependent variability of serum insulin concentrations: A comparison of eight assays. *Scand. J. Clin. Lab. Invest.* 77, 122–129. doi:10.1080/00365513.2016.1278260
- Tokarz, V. L., MacDonald, P. E., and Klip, A. (2018). The cell biology of systemic insulin function. *J. Cell. Biol.* 217, 2273–2289. doi:10.1083/jcb.201802095
- Topp, B., Promislow, K., Devries, G., Miura, R. M., and Finegood, D. T. (2000). A model of  $\beta$ -cell mass, insulin, and glucose kinetics: Pathways to diabetes. *J. Theor. Biol.* 206, 605–619. doi:10.1006/jtbi.2000.2150

- Tsilingiris, D., Koliaki, C., and Kokkinos, A. (2019). Remission of type 2 diabetes mellitus after bariatric surgery: Fact or fiction? *Int. J. Environ. Res. Public Health* 16, 3171. doi:10.3390/ijerph16173171
- Tsui, S. T., Yang, J., Zhang, X., Tatarian, T., Docimo, S., Spaniolas, K., et al. (2021). Health disparity in access to bariatric surgery. *Surg. Obes. Relat. Dis.* 17, 249–255. doi:10.1016/j.soard.2020.10.015
- Wallace, T. M., Levy, J. C., and Matthews, D. R. (2004). Use and abuse of HOMA modeling. *Diabetes Care* 27, 1487–1495. doi:10.2337/diacare.27.6.1487
- Wan, E. A., and van der Merwe, R. (2001). “The unscented kalman filter,” in *Kalman filtering and neural networks*. Editor S. Haykin (New York, USA: John Wiley & Sons), 221–280.
- Wang, G-F., Yan, Y-X., Xu, N., Yin, D., Hui, Y., Zhang, J. P., et al. (2015). Predictive factors of type 2 diabetes mellitus remission following bariatric surgery: A meta-analysis. *Obes. Surg.* 25, 199–208. doi:10.1007/s11695-014-1391-y
- Weiner, A., Cowell, A., McMahon, D. J., Tao, R., Zitsman, J., Oberfield, S. E., et al. (2020). The effects of adolescent laparoscopic adjustable gastric band and sleeve gastrectomy on markers of bone health and bone turnover. *Clin. Obes.* 10, e12411. doi:10.1111/cob.12411
- Weiskopf, N. G., Hripcsak, G., Swaminathan, S., and Weng, C. (2013). Defining and measuring completeness of electronic health records for secondary use. *J. Biomed. Inf.* 46, 830–836. doi:10.1016/j.jbi.2013.06.010
- Yan, W., Bai, R., Yan, M., and Song, M. (2017). Preoperative fasting plasma C-peptide levels as predictors of remission of type 2 diabetes mellitus after bariatric surgery: A systematic review and meta-analysis. *J. Investig. Surg.* 30, 383–393. doi:10.1080/08941939.2016.1259375
- Yeckel, C. W., Weiss, R., Dziura, J., Taksali, S. E., Dufour, S., Burgert, T. S., et al. (2004). Validation of insulin sensitivity indices from oral glucose tolerance test parameters in obese children and adolescents. *J. Clin. Endocrinol. Metab.* 89, 1096–1101. doi:10.1210/jc.2003-031503
- Yip, S., Plank, L. D., and Murphy, R. (2013). Gastric bypass and sleeve gastrectomy for type 2 diabetes: A systematic review and meta-analysis of outcomes. *Obes. Surg.* 23, 1994–2003. doi:10.1007/s11695-013-1030-z