

A GAUSSIAN PROCESS MODEL FOR INSULIN SECRETION RECONSTRUCTION WITH UNCERTAINTY QUANTIFICATION: APPLICATIONS IN CYSTIC FIBROSIS *

JUSTIN GARRISH[†], CHRISTINE L. CHAN[‡], DOUGLAS NYCHKA[†], AND CECILIA DINIZ
BEHN^{†‡}

Abstract. Investigation of biological systems often requires reconstruction of an unobservable continuous process from discrete time-series data sampled from a related process or processes. When the reconstructed process cannot be validated with experimental data, it is particularly important to quantify the uncertainty on the inferred process, and new methodologies are needed to support both the inference and uncertainty quantification. This work derives a novel statistical model that combines an established differential model of intra- and extravascular C-peptide dynamics with a Gaussian process model of insulin secretion rate (ISR) in order to provide clinical measures of beta-cell function with quantified uncertainty. These measures are computed from the ISR that is inferred from measured C-peptide data. The model is first validated using synthetic data, and then applied to oral glucose tolerance test (OGTT) data from youth participants with and without cystic fibrosis (CF). Because CF is characterized by scarring and fibrosis of the pancreas, impairment of beta-cell function, rather than reduced insulin sensitivity, is implicated in the early etiology of CF-related diabetes (CFRD). ISR-derived measures of beta-cell function show worsening beta-cell function from healthy control to CF to CFRD groups consistent with previous reports on dysglycemia in CF. However, the model additionally allows uncertainty in the data to be propagated to ISR and ISR-derived measures of beta-cell function. These results provide insight into uncertainty in ISR-derived measures of beta cell function, characterize interindividual variability in CFRD etiology, and provide novel metrics to quantify the pathogenesis of CFRD.

Key words. mathematical model, C-peptide, metabolism, uncertainty quantification, Gaussian process

AMS subject classifications. 92, 34, 62

1. Introduction. A common inverse problem that arises in mathematical physiology is to reconstruct an unobservable continuous process from discrete time-series data sampled from a related process or processes. Many methods have been proposed to solve this problem in different physiological contexts, however, current biological and medical research requires both inference of the process and quantification of the uncertainty on this inference, particularly when the underlying process cannot be measured experimentally. When model assumptions may be translated appropriately to (prior) distributions, a popular approach to solve such an inverse problem is the construction of a Bayesian hierarchical model [18]. By specifying conditional distributions between observed and unobserved processes given the required model parameters, and specifying all prior knowledge of parameter values as probability distributions, a posterior distribution can be inferred. This posterior distribution allows for estimation of the desired function(s) as well as quantification of uncertainty. In many cases, due to the parametric nature of the models used, this inference necessitates the random search of a multidimensional parameter space, for example using

*Funding for this study was provided by the National Science Foundation through grant DMS 1853511 and the National Institutes of Health through National Institute of Diabetes and Digestive and Kidney Diseases grants DK-094712-04, TR-000154 (Colorado Clinical Translational Sciences Institute), and UL1-TR-001082 (REDCap); and by Cystic Fibrosis Foundation grants CHAN16A0 and CHAN16GE0.

[†]Department of Applied Mathematics and Statistics, Colorado School of Mines

[‡]Department of Pediatrics, Section of Endocrinology, Children's Hospital Colorado, University of Colorado Anschutz Medical Campus

Markov chain Monte Carlo (MCMC) methods, e.g. in [9, 18]. However, by modeling unobserved phenomena as Gaussian processes and estimating empirically hyperparameters that are *a priori* unknown, it is possible to determine the resultant posterior distribution explicitly, and thus to obtain closed-form expressions for estimated curves and their confidence bounds.

An example of such an inverse problem is the reconstruction of an individual’s insulin secretion rate (ISR) profile from discrete samples of C-peptide concentration in blood plasma in the hours following a glucose stimulus. As ingested glucose enters the bloodstream and increases blood glucose concentrations, the beta cells of the pancreas release insulin, a hormone promoting glucose uptake throughout the body. Under healthy metabolic conditions, the pancreas releases sufficient insulin that plasma glucose returns back to basal levels in a relatively short time (e.g., 2 h). However, disruptions to metabolic function may result in higher glucose excursions and/or prolonged periods of elevated blood glucose. The time course of interactions between glucose and insulin may be formally evaluated with an oral glucose tolerance test (OGTT) and provide insight into the individual’s metabolic health. Metabolic disease may disrupt the glucose-insulin balance by reducing sensitivity to insulin over time as occurs in type-2 diabetes, or by causing damage to the beta cells and pancreas as occurs in type-1 diabetes and cystic fibrosis (CF).

CF is a life-limiting genetic disorder wherein a mutation in the CF transmembrane conductance regulator gene causes thick, dry mucous to restrict flow through narrow passages in the body, especially within the lungs and pancreas. Cystic fibrosis-related diabetes (CFRD) is the most common comorbidity associated with CF; up to 20% of adolescents and 30% to 50% of adults with CF develop CFRD [17], and incidence of CFRD is expected to rise as improved therapeutics for CF increase life expectancy for CF patients. Although insulin sensitivity is reduced in patients with CFRD, impaired beta-cell function is considered to be the primary contributor to early etiology of CFRD [2, 15, 19]. However, recent work has demonstrated that there are distinct changes in beta-cell function and insulin sensitivity across glucose tolerance stages in the progression to CFRD [19]. Loss of early insulin release leads to a pattern of elevated glucose levels at 140 min in an OGTT in CF patients well before they reach diagnostic criteria for CFRD, but even mild dysglycemia is associated with decreased pulmonary function in this population. With progressive loss of beta-cell function driven by myriad factors related to CF, insulin insufficiency leads to CFRD and is associated with increased mortality. [12].

Many methods have been proposed for assessing beta-cell function and include simple indices, deconvolution-based methods, and methods relating insulin secretion to glucose concentration [26, 3, 15, 14, 9]. The most sensitive methods for assessing beta-cell function consider ISR and can aid in monitoring the progression of metabolic dysregulation and disease. The rate of insulin secretion in response to a glucose challenge provides insight into an individual’s beta-cell function and aids in monitoring the progression of metabolic dysregulation and disease. The release of insulin, however, cannot be observed in real-time during a physiological oral test, thus motivating the need for inference of ISR from available data. C-peptide, a polypeptide produced in a 1-to-1 molar ratio with insulin and ultimately cleaved from the fully formed insulin molecule, is used as a surrogate measure of the insulin secreted by the pancreas because C-peptide, in contrast to insulin, is not cleared by the liver and is therefore a more accurate marker of the dynamics of secretion.

We propose a novel method for inferring insulin secretion rate (ISR) with uncertainty in a cohort of youth with and without CF. Our approach introduces a novel

statistical model that combines an established differential model of intra- and extravascular C-peptide dynamics with a Gaussian process model of ISR in order to provide clinical measures of beta-cell function with quantified uncertainty. To relate ISR to the measured values of C-peptide following an oral glucose stimulus during an OGTT, we employ the two-compartment model from [7]:

$$(1.1) \quad \begin{aligned} \dot{C}_1(t) &= -(k_1 + k_3)C_1(t) + k_2C_2(t) + S(t), & C_1(0) &= C_0, \\ \dot{C}_2(t) &= k_1C_1(t) - k_2C_2(t), & C_2(0) &= \frac{k_1}{k_2}C_0, \end{aligned}$$

where $C_1(t)$ represents the intravascular C-peptide concentration at time t (which can be observed at discrete times via blood sampling), $C_2(t)$ the extravascular C-peptide concentration, and $S(t)$ the ISR. The initial conditions are chosen so that the system is in equilibrium at time $t = 0$. For each participant, the kinetic parameters k_1 , k_2 , and k_3 are determined based on age and metabolic health. Van Cauter and colleagues identified standard parameters for C-peptide clearance accounting for obesity and age using a least-squares regression in a heterogeneous population of adults as described in [26]. They also showed that standard parameters rather than individually derived parameters could be used to model C-peptide clearance without a significant loss of accuracy. These parameters are taken as known in our model.

We demonstrate that, by modeling S as a Gaussian process, we can infer a closed form for $S(t)$ from discrete samples of C-peptide concentration taken during an OGTT. We apply this method to estimate several ISR-derived measures in order to characterize the impacts of CF on beta-cell function and glucose-insulin dynamics in our cohort.

The paper is structured as follows: [section 2](#) presents the materials and methods, including study design, mathematical and statistical modeling, and ISR-derived measures of beta-cell function and metabolic health; [section 3](#) details the results from analyzing synthetic data as well as clinical data from youth with and without CF; and lastly, [section 4](#) discusses the contributions of the paper to secretion rate modeling and the implications of the results for CF in youth.

2. Materials and Methods.

2.1. Study Participants. A cohort of 82 participants with and without CF aged 6 to 25 years were enrolled as part of a study of early glucose abnormalities in youth with CF. Appropriate consent and assent were obtained, and this study was approved by the Colorado Multiple Institutional Review Board. Participants without CF (healthy controls) were identified using recruitment flyers and emails at the University of Colorado Anschutz Medical Campus. Exclusion criteria for healthy controls included known diabetes or pre-diabetes, overweight or body mass index (BMI) greater than or equal to the 85th percentile according to percentile growth charts in youth, acute illness, history of chronic disease, or pregnancy. Participants with CF were recruited from pulmonary and diabetes clinics at Children’s Hospital Colorado. Inclusion criteria for participants with CF included diagnosis of CF by newborn screen, sweat chloride testing, or genetic testing. CF patients with known glucose abnormalities along the entire glycemic spectrum were included. Exclusion criteria for participants with CF included known type 1 or type 2 diabetes, use of medications affecting glucose metabolism (other than insulin) in the prior 3 months, hospitalization in the prior 6 weeks, or pregnancy. CFRD is defined as a 2h glucose concentration of >200 mg/dL or a fasting glucose concentration of >126 mg/dL observed during an OGTT [16]. Some participants with CFRD use exogenous insulin to

help regulate blood-glucose. For these insulin-dependent (ID) individuals, exogenous insulin was not administered for 24 hours prior to OGTT so measured glucose-insulin dynamics reflect physiological interactions. Among the 82 participants, 17 were classified as healthy controls, 51 as CF without CFRD (CF controls), and 12 as having CFRD. Four participants (1 healthy control, 2 CF control, and 1 CFRD) had missing data and were excluded from analysis.

2.2. Laboratory Procedures. Study visits took place in the Clinical Translational Research Center at Children’s Hospital Colorado between October 2015 and May 2018. Height, weight, and vital signs were obtained. A physical exam, including pubertal staging using the method of Tanner, was completed by a single pediatric endocrinologist. Participants arrived to the outpatient research center between 8 a.m. and 10 a.m. after a minimum of 8 hours of fasting. All participants underwent oral OGTT. Glucola was administered at a dose of 1.75 g/kg (maximum 75 g). As the study involved youth participants, a reduced 3-hour schedule was chosen to minimize the strain and physical burden of the OGTT on the participants [5]. Measurements of glucose, insulin, and C-peptide were taken at -10, 0, 10, 20, 30, 60, 90, 120, 150, and 180 minutes with time 0 corresponding to the administration time of Glucola. Time -10 corresponds to 10 mins prior to Glucola administration and was used to establish basal values of glucose, insulin, and C-peptide. The protocol includes more frequent sampling in the first half hour following Glucola administration when measured metabolites are expected to change most quickly.

2.3. Data and Process Models. The differential system (1.1) relating ISR to the measured C-peptide values can be rewritten in matrix-vector form as

$$(2.1) \quad \begin{bmatrix} \dot{C}_1(t) \\ \dot{C}_2(t) \end{bmatrix} = \begin{bmatrix} -k_1 - k_3 & k_2 \\ k_1 & -k_2 \end{bmatrix} \begin{bmatrix} C_1(t) \\ C_2(t) \end{bmatrix} + \begin{bmatrix} S(t) \\ 0 \end{bmatrix}.$$

Equation (2.1) demonstrates that (1.1) is linear and non-homogeneous, with forcing provided by $S(t)$. Suppose the matrix of coefficients in (2.1) has eigenvalue-eigenvector pairs $\{\lambda_1, [u_1, u_2]^T\}$ and $\{\lambda_2, [v_1, v_2]^T\}$, and define above basal C-peptide and ISR c and s , respectively, as

$$\begin{aligned} c(t) &:= C_1(t) - C_0, \text{ and} \\ s(t) &:= S(t) - S(0) = S(t) - k_3 C_0. \end{aligned}$$

With this change of variables, the integral equation form of (2.1) reduces to

$$(2.2) \quad c(t) = \frac{1}{u_1 v_2 - v_1 u_2} \left[u_1 v_2 e^{\lambda_1 t} \int_0^t e^{-\lambda_1 \tau} s(\tau) d\tau - v_1 u_2 e^{\lambda_2 t} \int_0^t e^{-\lambda_2 \tau} s(\tau) d\tau \right]$$

in terms of c and s . Thus, there is a linear transformation $\mathcal{W}$ such that $c = \mathcal{W}s$.

Let $\{\tau_0, \dots, \tau_m\}$ be the sampling times and $\{t_1, \dots, t_n\}$ be the times at which we estimate ISR with $\{\tau_1, \dots, \tau_m\} \subset \{t_1, \dots, t_n\}$. Consider $\mathbf{c} = [c(\tau_1) \cdots c(\tau_m)]^T \in \mathbb{R}^m$ and $\mathbf{s} = [s(t_1) \cdots s(t_n)]^T \in \mathbb{R}^n$, where $n > m$. Then, $\mathbf{c}$ and $\mathbf{s}$ must satisfy

$$\mathbf{c} = \mathbf{W}\mathbf{s} + \mathbf{e},$$

where the $m \times n$ matrix $\mathbf{W}$ discretizes the continuous linear transformation $\mathcal{W}$, and $\mathbf{e}$ represents the numerical error in estimating the integrals from (2.1) using a quadrature rule which tends to $\mathbf{0}$ as $n \rightarrow \infty$. We are able to make n much larger than m : the

OGTT protocol has $m = 8$ sampling points, and we take $n = 180$. Therefore, these numerical errors are negligible relative to the observational errors noted below, and so we will not consider the influence of $\mathbf{e}$ in subsequent analysis.

The measured above-basal C-peptide values, $\mathbf{z}$, are taken with observation errors $\boldsymbol{\varepsilon}$ which are assumed to be independent and normally distributed:

$$(2.3) \quad \mathbf{z} = W\mathbf{s} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \text{MN}(\mathbf{0}, \Sigma_C),$$

where Σ_C is diagonal, and $\text{MN}(\boldsymbol{\mu}, \Sigma)$ denotes a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix Σ . In addition, we employ a mixed-effects model for the above-basal ISR,

$$(2.4) \quad s_i := s(t_i) = \alpha_1 + \alpha_2 t_i + y(t_i),$$

where α_1 and α_2 are fixed parameters and y is the realization of a zero-mean Gaussian process (GP) with Matérn covariance function [8]

$$(2.5) \quad \text{Cov}(y(t), y(\tau)) = \sigma^2 k_\theta(t, \tau) = \sigma^2 \left(1 + \frac{|t - \tau|}{\theta} \right) \exp \left(-\frac{|t - \tau|}{\theta} \right).$$

The Matérn family of covariances is a common choice in Gaussian process modeling and offers varying degrees of smoothness with a small number of (hyper)parameters left to estimate [23]. In general, a Matérn process with smoothness ν has m mean-square derivatives for $m < \nu$. The function (2.5) is obtained by setting the smoothness to $\nu = 1.5$, so that any given sample path is C^1 -smooth. This is a reasonable assumption given the process that we are modeling, and this is the minimal smoothness assumption necessary to enforce the steady-state boundary conditions on our ISR estimate.

In (2.5), $y(t)$ denotes the value of the GP at time t , σ^2 represents the variance of the GP, and θ represents the time-range of the covariance. This covariance function models a GP with sample paths that are continuous with continuous first derivative. Since y is a GP, the random vector

$$\mathbf{y} := [y(t_1) \cdots y(t_n)]^T$$

has a multivariate normal distribution with covariance matrix $\sigma^2 K_\theta$, inherited from the function $\sigma^2 k_\theta$ as

$$(K_\theta)_{ij} = k_\theta(t_i, t_j).$$

The distributions for the GP and its evaluations on the n -node evaluation grid may then be summarized concisely as

$$(2.6) \quad \begin{aligned} y &\sim \text{GP}(0, \sigma^2 k_\theta), \text{ and} \\ \mathbf{y} &\sim \text{MN}(\mathbf{0}, \sigma^2 K_\theta). \end{aligned}$$

Defining matrix X and vector $\boldsymbol{\alpha}$ as

$$(2.7) \quad X = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} \text{ and } \boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix},$$

we may rewrite (2.4) as

$$(2.8) \quad \mathbf{s} = X\boldsymbol{\alpha} + \mathbf{y},$$

so that (2.3) becomes

$$(2.9) \quad \mathbf{z} = WX\boldsymbol{\alpha} + W\mathbf{y} + \boldsymbol{\varepsilon}.$$

The final assumption in our model is that the GP y and the measurement errors are independent. Setting $V = WX$ and $\boldsymbol{\eta} = W\mathbf{y} + \boldsymbol{\varepsilon}$, we rewrite (2.9) as

$$\mathbf{z} = V\boldsymbol{\alpha} + \boldsymbol{\eta},$$

If we denote the covariance matrix for $\boldsymbol{\eta}$ by H , then using linear statistics,

$$\mathbf{z} \sim \text{MN}(V\boldsymbol{\alpha}, H),$$

where

$$H = \text{Cov}(W\mathbf{y} + \boldsymbol{\varepsilon}) = \sigma^2 WK_\theta W^T + \Sigma_C.$$

2.4. Steady-State Condition. Participants undergo a minimum 8 h fast prior to the OGTT, so the plasma glucose, insulin, and C-peptide concentrations are assumed to be at steady state at the beginning of the protocol. To ensure that the system (1.1) is in steady state when $t = 0$, we enforce the initial condition $s(0) = s'(0) = 0$, i.e. that the above-basal ISR and its derivative are 0 prior to the OGTT. From (2.4), we have that $y_0 := y(0) = -\alpha_1$, and $y'_0 := y'(0) = -\alpha_2$. Thus, we condition our inference on the assumption that $\mathbf{y}_0 := [y(0) \ y'(0)]^T = [-\alpha_1, \alpha_2]^T = -\boldsymbol{\alpha}$.

Because we model y as a GP, we have the following multivariate normal distribution for the vector $[\mathbf{y}_0^T \ \mathbf{y}^T]^T$:

$$(2.10) \quad \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y} \end{bmatrix} \sim \text{MN} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \sigma^2 \begin{bmatrix} K_{00} & K_{01} \\ K_{10} & K_{11} \end{bmatrix} \right),$$

where the zero vectors are assumed to be of appropriate dimension, and sub-matrices K_{00} , K_{01} , K_{10} , K_{11} are assembled using equation (2.5) and equations (A.1), (A.3), (A.4), and (A.5) found in the appendix. In the K_{ij} sub-matrices, the subscripts t and τ are used to denote partial differentiation with respect to the first argument and second argument, respectively. These sub-matrices are defined as

$$K_{00} = \begin{bmatrix} k_\theta(0, 0) & (k_\theta)_t(0, 0) \\ (k_\theta)_\tau(0, 0) & (k_\theta)_{\tau t}(0, 0) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\theta^2} \end{bmatrix},$$

$$K_{01} = \begin{bmatrix} k_\theta(0, t_1) & \cdots & k_\theta(0, t_n) \\ (k_\theta)_t(0, t_1) & \cdots & (k_\theta)_t(0, t_n) \end{bmatrix},$$

$$K_{10} = \begin{bmatrix} k_\theta(t_1, 0) & (k_\theta)_\tau(t_1, 0) \\ \vdots & \vdots \\ k_\theta(t_n, 0) & (k_\theta)_\tau(t_n, 0) \end{bmatrix} = K_{01}^T,$$

$$K_{11} = K_\theta.$$

Given that $\mathbf{y}_0 = -\boldsymbol{\alpha}$, we obtain the following conditional expectation and covariance for the random vector $\mathbf{y}$:

$$\begin{aligned} \boldsymbol{\mu}^* &= \text{E}(\mathbf{y} | \mathbf{y}_0 = -\boldsymbol{\alpha}) = -K_{10}K_{00}^{-1}\boldsymbol{\alpha}, \text{ and} \\ \sigma^2 K^* &= \text{Cov}(\mathbf{y} | \mathbf{y}_0 = -\boldsymbol{\alpha}) = \sigma^2 (K_\theta - K_{01}K_{00}^{-1}K_{10}). \end{aligned}$$

249 We also condition the covariance of $\boldsymbol{\eta}$:

$$250 \quad H^* = \text{Cov}(\boldsymbol{\eta} | \mathbf{y}_0 = -\boldsymbol{\alpha}) = \sigma^2 W K^* W^T + \Sigma_C.$$

251 **2.5. Hyperparameter Estimation.** Denote by $\ell(\sigma^2, \theta, \boldsymbol{\alpha}; \mathbf{z})$ the log-likelihood
 252 of our hyperparameters given an individual's measured C-peptide, $\mathbf{z}$. The Gaussian
 253 and other model assumptions allow us to write the individual log-likelihood as

$$254 \quad (2.11) \quad \ell(\sigma^2, \theta, \boldsymbol{\alpha}; \mathbf{z}) = -\frac{1}{2} \left[(\mathbf{z} - V\boldsymbol{\alpha})^T H^{*-1} (\mathbf{z} - V\boldsymbol{\alpha}) - m \ln(2\pi) - \ln(\det(H^*)) \right].$$

255 For any given σ^2 and θ , we can apply generalized least squares (GLS) to estimate
 256 $\boldsymbol{\alpha}$ for the individual. Setting $V = WX$ and $\boldsymbol{\eta} = W\mathbf{y} + \boldsymbol{\varepsilon}$, we rewrite (2.9) as

$$257 \quad \mathbf{z} = V\boldsymbol{\alpha} + \boldsymbol{\eta}.$$

258 Then, the GLS estimate of $\boldsymbol{\alpha}$ is given by

$$259 \quad \hat{\boldsymbol{\alpha}}_{\sigma^2, \theta} = (V^T H^{*-1} V)^{-1} V^T H^{*-1} \mathbf{z},$$

260 where the subscripts emphasize dependence on the hyperparameters (via dependence
 261 on H). We use $\hat{\boldsymbol{\alpha}}_{\sigma^2, \theta}$ in all subsequent calculations depending on $\boldsymbol{\alpha}$.

262 Following an empirical Bayes approach, we borrow strength across the C-peptide
 263 measurements of all participants to estimate cohort values for hyperparameters σ^2
 264 and θ by maximum likelihood. If p participants' data are analyzed, and the j -th par-
 265 ticipant has above-basal C-peptide values $\mathbf{z}_j$, then the cohort log-likelihood $\ell_c(\sigma^2, \theta)$
 266 is given by

$$267 \quad (2.12) \quad \ell_c(\sigma^2, \theta) = \sum_{j=1}^p \ell(\sigma^2, \theta; (\hat{\boldsymbol{\alpha}}_{j, (\sigma^2, \theta)}, \mathbf{z}_j).$$

268 We then estimate hyperparameters for the cohort as

$$269 \quad (2.13) \quad (\hat{\sigma}^2, \hat{\theta}) = \arg \max_{\sigma^2, \theta} \ell_c(\sigma^2, \theta).$$

270 Although the cohort log-likelihood is nonlinear, a grid search is reasonable to de-
 271 termine $\hat{\sigma}^2$ and $\hat{\theta}$ based on the moderate size of this problem and also allows for
 272 empirically verifying a unique maximum.

273 **2.6. ISR Reconstruction and Uncertainty Quantification.** With hyper-
 274 parameters σ^2 and θ fixed to empirically determined values $\hat{\sigma}^2$ and $\hat{\theta}$, and with $\boldsymbol{\alpha}$
 275 fixed to the GLS estimate $\hat{\boldsymbol{\alpha}}$, we may approximate the conditional expectation of
 276 $[\mathbf{y} | \sigma^2, \theta, \boldsymbol{\alpha}, \mathbf{y}_0]$ by substituting these estimates of the statistical parameters into the
 277 conditional distribution to obtain our estimate of $\mathbf{y}$:

$$278 \quad (2.14) \quad \hat{\mathbf{y}} = \boldsymbol{\mu}^* + (W K^*)^T H^{*-1} (\mathbf{z} - V\hat{\boldsymbol{\alpha}} - W\boldsymbol{\mu}^*).$$

279 We subsequently reconstruct the ISR profile $\mathbf{s}$ as

$$280 \quad (2.15) \quad \hat{\mathbf{s}} = X\hat{\boldsymbol{\alpha}} + \hat{\mathbf{y}},$$

281 which is the best linear unbiased estimator of above-basal ISR with known σ^2 and θ .
 282 Adding the steady-state secretion rate, $k_3 C_0$, to each component of $\hat{\mathbf{s}}$ yields the total
 283 ISR approximation, $\hat{\mathbf{S}}$.

Consider the $n \times m$ matrix A whose i -th column, $\mathbf{a}_i$, is given by

$$(2.16) \quad \mathbf{a}_i = \boldsymbol{\mu}^* + (WK^*)^T H^{*-1} (\mathbf{e}_i - V\boldsymbol{\alpha} - W\boldsymbol{\mu}^*),$$

where $\mathbf{e}_i$ is the i -th standard basis vector in $\mathbb{R}^m$. To quantify uncertainty in our inference, we can compute the standard error from the covariance of $\hat{\mathbf{S}}$, given by

$$(2.17) \quad \sigma^2 \hat{K} = \sigma^2 (K^* - (AWK^*)^T H^{*-1} AWK^*),$$

so that we obtain the standard error vector from the square roots of the diagonal elements of $\sigma^2 \hat{K}$,

$$\mathbf{SE} = \begin{bmatrix} \sigma \sqrt{\hat{K}_{11}} \\ \vdots \\ \sigma \sqrt{\hat{K}_{nn}} \end{bmatrix}.$$

The pointwise 95% confidence envelope for $\hat{\mathbf{S}}$ is then given by $\hat{\mathbf{S}} \pm 1.96 \cdot \mathbf{SE}$. The Bonferroni correction with 95% confidence and $m = 8$ degrees of freedom (corresponding to the number of C-peptide measurements during the OGTT) yields conservative upper and lower bounds for the 95% confidence envelope with uniform coverage of ISR profiles comprising the envelope as $\hat{\mathbf{S}} \pm 2.73 \cdot \mathbf{SE}$.

2.7. Application of Model to Synthetic and Clinical Data. Since the true ISR profile during an OGTT cannot be measured experimentally, we validated this method on simulated data for which the ISR was known by verifying that we recovered a simulated ISR profile with good confidence from simulated C-peptide data. A continuous ISR profile was simulated by inferring the ISR from a healthy participant's OGTT data. We used this known ISR to simulate 3-hour OGTT C-peptide data by applying the differential model (1.1) and sampling at time points determined by the clinical sampling schedule. We computed the pointwise confidence bounds with a Bonferroni correction as described in subsection 2.6.

The ISR reconstruction approach was also applied to the C-peptide data collected from the cohort of control, CF control, and CFRD participants. For the inferred ISR we computed the pointwise confidence bounds with a Bonferroni correction as described in subsection 2.6.

2.8. ISR-Based Measures of Beta-Cell Function. A functional reconstruction of ISR provides several key measures to assess different features of beta-cell function. We analyze the ISR area under the curve (AUC) in the first 30 min of the OGTT; the ISR AUC over the full 3 h OGTT; ISR time-to-peak (TTP); and the slope α_2 of the linear component of the mixed-effects model (2.8). All AUCs were computed numerically by composite trapezoid rule. The numerical approximation error will be small due to the flexibility in choosing a large number of node points (n) for estimating the ISR curves.

Because the covariance provides a complete description of uncertainty (2.17), linear statistics enable us to propagate uncertainty in the ISR prediction through to indices of beta-cell function. For example, if $\mathbf{w} \in \mathbb{R}^n$ is a vector of weights associated with composite trapezoid rule, the covariance of the ISR AUC estimate, $\mathbf{w}^T \mathbf{s}$, is given by

$$(2.18) \quad \text{Cov}(\mathbf{w}^T \mathbf{s}) = \mathbf{w}^T \text{Cov}(\mathbf{s}) \mathbf{w} = \sigma^2 \mathbf{w}^T \hat{K} \mathbf{w}.$$

Thus, the covariance of the ISR AUC estimate can be directly related to the covariance of ISR itself.

2.9. Statistical Analyses. We analyze differences between measures across groups using the Kruskal-Wallis rank-sum test due to factors such as non-normality of computed measures, non-uniformity of sample sizes, and relatively small sample sizes for the healthy control and CFRD groups. In cases where differences between group medians are observed with 95% confidence, post hoc analyses are performed with the pairwise Wilcoxon rank-sum test. The p -values are reported, and comparisons with $p < 0.05$ are considered statistically significant.

3. Results.

3.1. Cohort Hyperparameters. The number of polynomial basis function used in the mixed-effects model (2.4) was determined by assessing the log-likelihood surfaces with respect to the identifiability of the model hyperparameters. We performed the cohort hyperparameter estimation with mixed-effects models of the form

$$s(t) = \sum_{j=1}^d \alpha_j t^{j-1} + y(t),$$

for varying values of d ($d = 1, 2, 3, 4$). We found that the model with $d = 2$ yielded a log-likelihood surface that best identified the hyperparameters and confidence sets.

By computing the cohort log-likelihood (2.12) on a fine grid of values for hyperparameters θ and σ^2 , we estimate that maximum likelihood is attained at $\hat{\theta} = 17$ min and $\hat{\sigma}^2 = 7400$ pmol²/L² · min². The optimal cohort values with 95% and 99% confidence sets are shown in Figure 3.1.

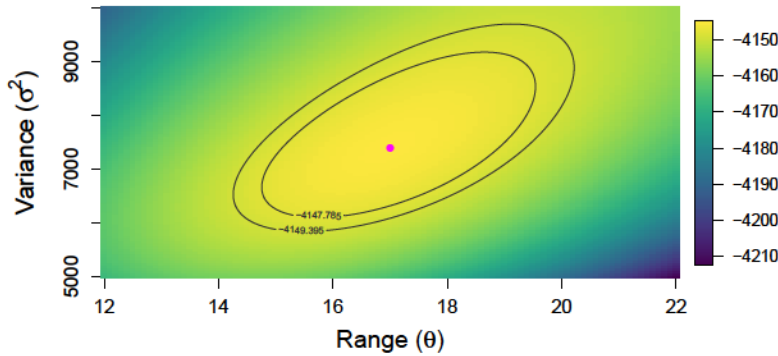


FIG. 3.1. Log-likelihood values for combinations of hyperparameters θ and σ^2 , with maximum likelihood achieved for $\theta = 17$ and $\sigma^2 = 7400$. The contours show 95% and 99% confidence sets.

3.2. Analysis of Simulated Data. We used the simulated ISR profile and corresponding simulated OGTT C-peptide data shown in Figure 3.2 to validate the model. Applying our method to the simulated noisy C-peptide observations, we inferred the ISR profile and placed 95% confidence bounds on the inferred curve (Figure 3.3). We found close agreement between the true and inferred ISR for this example with synthetic data, and the true ISR was within the 95% confidence bounds of the inferred ISR curve. Note that the 95% confidence bounds are computed pointwise

and are therefore tightest near the points where data was collected. Due to the forward transformation of ISR to C-peptide concentration and inherent averaging in the reconstruction of ISR, the time points corresponding to minimal uncertainty will not be mapped directly to OGTT sampling times.

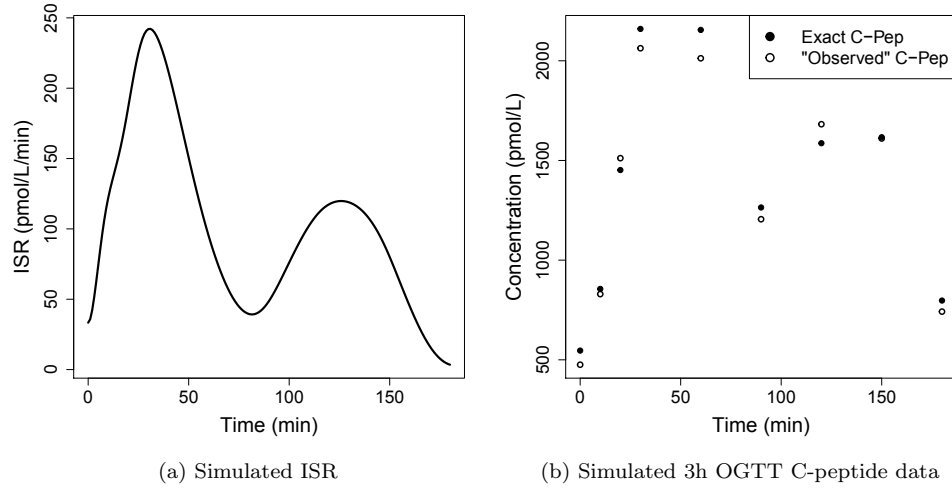


FIG. 3.2. (a) ISR profile simulated for a clinical 3h OGTT. (b) Simulated exact C-peptide data (filled) and noisy C-peptide data (open) sampled according to clinical OGTT schedule.

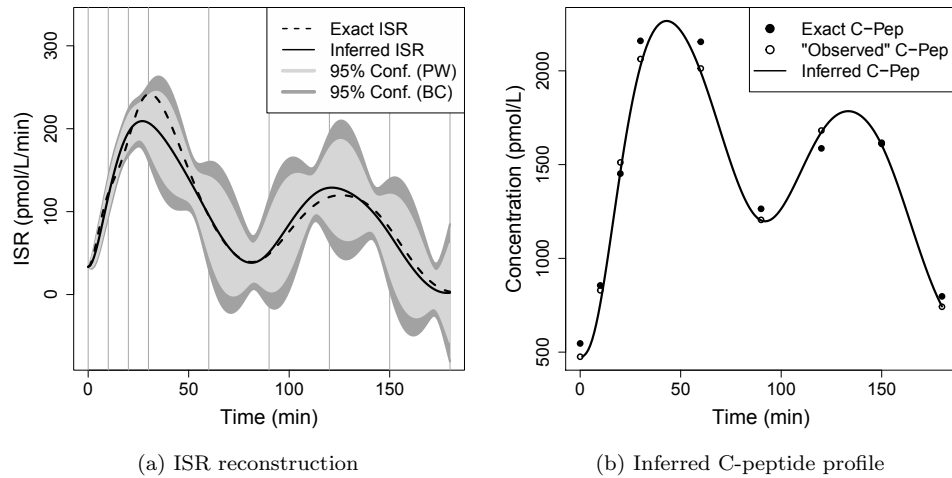


FIG. 3.3. (a) ISR profile reconstructed using noisy data representing observed C-peptide, with 95% confidence bounds containing the simulated ("exact") ISR curve (dashed). Vertical lines represent sampling times. (b) C-peptide profile modeled using the reconstructed ISR using the C-peptide minimal model is consistent with the simulated C-peptide data.

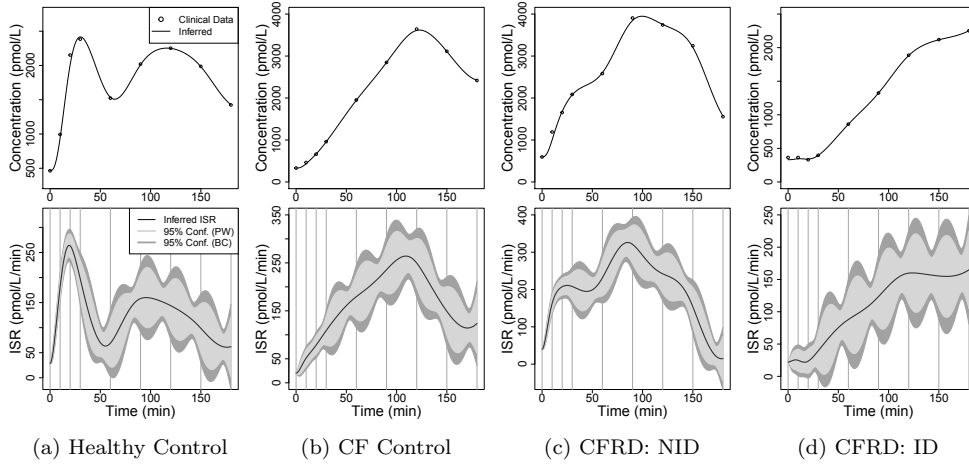


FIG. 3.4. (Top) C-peptide data collected during OGTT with C-peptide profiles estimated from inferred ISR. (Bottom) ISR profiles reconstructed from OGTT data with 95% confidence sets. Vertical lines represent OGTT sampling times. From left to right, sample reconstructions for a (a) healthy control participant, (b) CF control participant, (c) non-insulin-dependent CFRD participant (CFRD: NID), and (d) insulin-dependent CFRD participant (CFRD: ID).

3.3. Analysis of Clinical Data. We present the results from applying our method to OGTT C-peptide data from youth participants with and without CF. We include some representative examples of ISR reconstruction and discuss ISR-derived measures of beta cell function for the full cohort.

3.3.1. ISR Reconstruction. OGTT data and corresponding inferences for ISR and C-peptide are shown in Figure 3.4 for representative participants across the three groups (healthy control, CF control, and CFRD). Within the CFRD group, we include results from two participants: one non-insulin-dependent (NID) individual and one insulin-dependent (ID) individual. The measured C-peptide profiles for these representative participants demonstrate a range of features including multiple peaks (Figure 3.4A), delayed time to peak (Figure 3.4B, C), and a gradual upward drift for the duration of the OGTT (Figure 3.4D). These features are recapitulated by the inferred ISR profiles.

3.3.2. ISR-Based Measures of Beta-Cell Function. In Table 3.1, we present the non-parametric analysis of model-based measures of beta-cell function based on the inferred ISR profile including 30 min ISR AUC, 3 h ISR AUC, ISR TTP, and trendline slope (α_2). Median values increase for each of 3 h ISR AUC, ISR TTP, and trendline slope, α_2 , from the healthy control values to the CF control values to the CFRD values. Inversely, the median values for the 30-min ISR AUC decrease from healthy controls to CF controls to the CFRD group. As statistically significant differences were observed across the groups for each measure considered, post hoc results were computed (Table 3.2). The post hoc tests showed statistically significant differences with at least 95% confidence across all pairs of groups for 3-h ISR AUC and ISR TTP. For 30-min AUC, a statistically significant difference was observed between the two CF groups and the healthy controls, and a statistically significant difference was observed in the slope α_2 between the CFRD group and the two control groups.

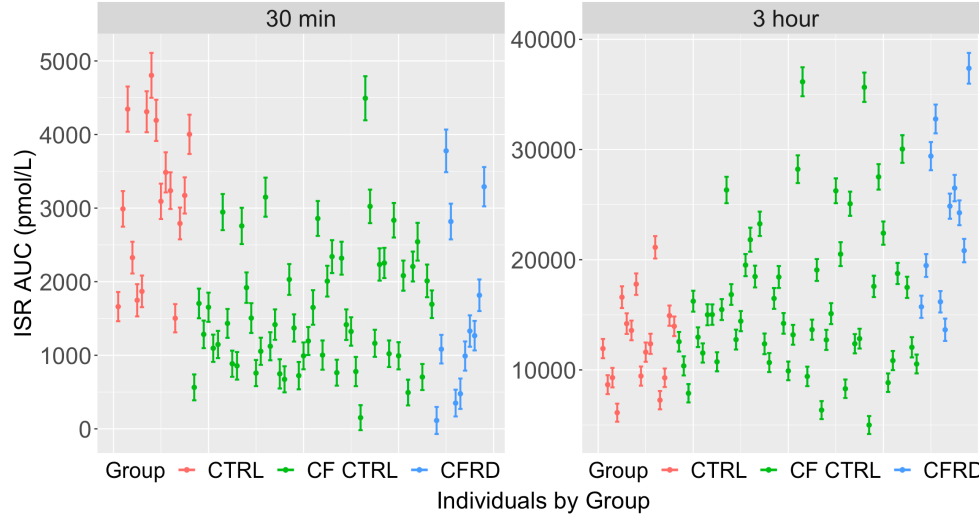


FIG. 3.5. *ISR AUC estimates with confidence bounds. (Left) AUC computed over first 30 min. (Right) AUC computed over duration of OGTT (3 hours).*

Figure 3.5 depicts ISR AUC for the first 30 mins of the OGTT and ISR AUC for the full 3 hour OGTT with 95% confidence intervals for each participant. In general, the amount of insulin secreted within the first 30 mins of the OGTT was suppressed in participants with CF. By contrast, the AUC inferred for the entire OGTT was highest in the CFRD group and lowest in the control group. This summary of individual data highlights overall trends of as well as the range of interindividual differences in beta-cell function in these groups.

TABLE 3.1

ISR-based results from non-parametric (Kruskal-Wallis) test. Quantities analyzed are ISR area under the curve (AUC), ISR time to peak (TTP), and slope of the trendline (α_2). Median values are shown for each group, and a statistically significant difference is observed between medians for each quantity analyzed with at least 95% confidence.

Quantity	CTRL	CF CTRL	CFRD	H-Statistic	p-Value
Sample Size	16	51	11	—	—
30min ISR AUC (pmol/L)	3132.078	1415.844	1266.425	19.93	0.0005
3h ISR AUC (pmol/L)	12142.89	14998.13	24190.61	15.38	0.0005
ISR TTP (min)	26.5	72.0	100.0	34.68	<0.0001
α_2 (pmol/L/min ²)	0.282	0.363	0.858	8.54	0.0139

TABLE 3.2

ISR-based results from pairwise Wilcoxon analysis.

Quantity	CTRL & CF CTRL	CF CTRL & CFRD	CTRL & CFRD
30min ISR AUC	<0.0001	0.6450	0.0049
3h ISR AUC	0.0336	0.0057	0.0002
ISR TTP	<0.0001	0.0048	<0.0001
α_2	0.2086	0.0319	0.0098

3.3.3. Sensitivity of Results to Kinetic Parameters. As described in the Methods, the kinetic parameters k_1 , k_2 , and k_3 were specified using the equations described by Van Cauter and colleagues in [26] and taken as known. To determine the sensitivity of the model estimates to the kinetic parameters k_1 , k_2 , and k_3 , we simulated ISR and computed 30 min ISR AUC and 3 h ISR AUC for C-peptide data from our cohort with a distribution of kinetic parameters. This distribution was determined by combinations of health statuses considered in the Van Cauter equations with ages ranging from 6 to 50 years. We found that the 30 min ISR AUC varied by less than 5% from the mean and the 3 h ISR AUC varied by less than 9% from the mean.

4. Discussion.

4.1. Summary of Results. We have demonstrated that our method successfully reconstructs a participant’s ISR profile from C-peptide OGTT data, with quantifiable uncertainty that can be propagated through to computed metrics of beta-cell function that depend on ISR. The ability to pool all participant data together to estimate hyperparameter values using a simple maximum likelihood approach offers value to researchers applying this method by allowing for increased predictive power despite sparse sampling schedules. The 30 min ISR AUC and TTP values calculated from reconstructed ISR profiles demonstrate a suppressed early insulin response among the CF control group compared to the control group. This suppressed early insulin response occurs in conjunction with an increase in the overall insulin released as assessed by the 3 h ISR AUC. Both of these effects are amplified with the progression to CFRD and suggest deterioration of beta cell function likely caused by the scarring and fibrosis of the pancreas that occur in CF. These findings are consistent with previous results on metabolic dysregulation in people with CF [4, 22, 25] and provide support for the presented statistical model as a clinically relevant tool for ISR inference. In particular, the simple ISR AUC offers an easily calculable index of beta-cell function that, in conjunction with other indices, may be useful in monitoring metabolic health for individuals with (or at risk of developing) metabolic disorders, including those originating from CF.

The differences in the individual 30-min ISR AUC and 3-hour ISR AUC (Figure 3.5) within CF and CFRD groups illustrate the individual and often insidious nature of CFRD onset and progression. For instance, an individual with CF that does not meet the OGTT-determined CFRD criteria may still exhibit metabolic dysfunction via delayed insulin action, which may be a manifestation of a suppressed first-phase insulin response, reduced sensitivity to insulin, or a combination of the two.

4.2. Motivation for Statistical Model. The covariance structure is a key component of the model presented, as it places the estimates on a firm mathematical foundation with avenues to stable and efficient computation even given sparse sampling and ill-conditioned systems that are common in ISR reconstruction and other applications in mathematical physiology. As the choice of covariance function used to model ISR determines the function space for our possible ISR inferences, additional physiologic information may be built in to improve ISR estimates and confidence bounds. Pulsatile insulin secretion [20, 24], for example, is a known phenomenon that may inform the construction of a customized covariance function for ISR inference.

Several methods in the literature use deconvolution to infer ISR from C-peptide data at varying levels of sparsity [7, 9, 10, 15, 20]. In [3, 5], deconvolution is applied

instead to glucose data in order to estimate the derivative of glucose during the OGTT. While many of the existing deconvolution methods are designed for metabolic experiments with more frequent sampling, [3] and [9] may be applied to OGTT data with a sparse sampling schedule. A common strategy in this family of methods is a penalized least squares approach, where fidelity to data is optimized against some penalized term(s), often representing roughness or other traits undesirable in the estimate. Roughness is commonly quantified via the norm of a finite-difference matrix [6, 9], controlling some derivative of the inferred profile. While allowing for control over desired features of the reconstruction, these methods typically require large-scale random search, e.g. MCMC, in order to achieve physiologically realistic profiles and to also quantify uncertainty.

Our approach avoids such large-scale Monte Carlo searches by invoking the linear properties of Gaussian processes and writing all distribution information explicitly. Nonetheless, the method maintains predictive power stemming from the choice of covariance function used to model the Gaussian process and the calculation of cohort-wide hyperparameters.

The mixed-effects model (2.4) and choice of covariance in (2.5) offer several advantages to ISR reconstruction from sparse C-peptide OGTT data. First, the mixed-effects nature has the flexibility to include additional covariates that support estimation of ISR and stable numerical computation, e.g. in (2.14), involving potentially ill-conditioned matrices. Additionally, the Matérn covariance function chosen guarantees that the estimated ISR profile will be sufficiently smooth to enforce the initial steady-state condition. Numerical stability is supported by regularization of our estimate [6], and we note that this arises naturally in the form of the conditional distribution. There is an equivalence between the empirical Bayesian problem and the unconstrained variational problem

$$(4.1) \quad \min_{\mathbf{y} \in \mathbb{R}^n} (\mathbf{z} - W(X\hat{\alpha} - \mathbf{y}))^T \Sigma_C^{-1} (\mathbf{z} - W(X\hat{\alpha} - \mathbf{y})) + \sigma^2 \mathbf{y}^T K^{-1} \mathbf{y}.$$

In (4.1), the second term is interpreted as a penalty with σ^2 controlling its weight in the optimization. This variational formulation alone is not rich enough to provide estimates of σ^2 and θ and is in contrast to a likelihood or Bayesian approach.

4.3. Kinetic Parameters and Time Scales. The kinetic parameters k_1 , k_2 , and k_3 specify time constants ($1/\lambda_1$ and $1/\lambda_2$) for the model (2.1) on the order of 5 and 35 min. These time constants reflect properties of C-peptide clearance and exchange rates in participants based on biometrics, and they are determined independently of the OGTT data. By contrast, the time constant θ that appears in the Gaussian process reflects the model’s ability to resolve the ISR dynamics given the sampling schedule and variation in observed values. We pooled the data from the entire cohort of participants to arrive at our estimate $\hat{\theta}$ on the order of 17 min. With this time-range, values of ISR located more than 29 min apart have a correlation of 0.5 or less. Data sets with different data collection protocols or participant populations may yield different hyperparameter values, affecting the range of time points with strong autocorrelation. For example, with more frequent sampling, the time-range of the GP may be resolvable to the fast time-scale observed in the differential C-peptide model. The credible intervals of the ISR estimates will depend on the interactions between these contributing factors.

4.4. Limitations of Current Modeling Approach. There are several limitations to this modeling approach. This method allows for estimation of aphysiological

negative values of ISR and the 95% confidence envelope. From the overall cohort of 82 participants, seven individuals' (representing all three groups) ISR profiles were estimated to have negative values. However, these negative excursions were typically small and brief, and therefore, minimally affected measures of beta-cell function that depend on the ISR profile. This issue has been observed in other methods of ISR reconstruction, and has been addressed by adding an additional penalty term to the cost function in the governing optimization problem [9] or by constraining the optimization to preclude non-negative values [10]. Alternatively, in future work we can modify our modeling structure to build non-negativity into the process level of the Bayesian hierarchical model rather than add an additional penalty term to the objective function being optimized. This will allow numerical implementation to remain cost-efficient and avoid large-scale simulation techniques such as MCMC while still increasing the model's predictive power.

In addition, the kinetic parameters k_1 , k_2 , and k_3 are taken as known in our model based on results established in adults with and without type 2 diabetes [26]. Our numerical experiments with the kinetic parameters indicate that the method and resulting estimates of 30 min and 3 h ISR AUC are much more sensitive to the C-peptide data than to kinetic parameters derived from biometric input. Although this suggests that uncertainty in the data affects estimated ISR more strongly than uncertainty in the kinetic parameters, the biometric dependence of kinetic parameters established in [26] has not been validated in pediatric populations or in patients with cystic fibrosis. Bayesian approaches to estimate these parameters more generally have been proposed [21]. However, improved estimates for these parameters are needed in populations with diverse genetic backgrounds, body types, ages, and disease status. Moreover, future approaches considering uncertainty in kinetic parameters could propagate this additional source of uncertainty to measures such as 30 min and 3 h ISR AUC, thereby further improving confidence in estimates of these quantities.

ISR is known to depend on glucose concentration and its derivative, and this relationship has been exploited in other models of ISR [3, 15, 24, 13, 1]. Our current model structure does not include glucose data. However, the flexibility of the mixed-effects structure will allow future incorporation of glucose and its derivative into the model in such a way that uncertainty in the data is propagated to all inferred continuous profiles and model-derived indices. One possible way to add glucose to our current model is to use a sigmoidal dependence of ISR on glucose, as proposed in [13, 24, 1], in the fixed component of ISR. Modifying the model to include glucose or other covariates will allow for additional dynamics-motivated metrics to be estimated that can provide a more nuanced picture of individual health and the landscapes of heterogeneous metabolic diseases such as CFRD.

4.5. Implications of Results for Clinical Practice. Most methods for reconstructing ISR have been developed in the context of an IVGTT [11]. However, OGTTs are the most commonly used protocols for assessing glycemic control in a clinical setting, and annual OGTTs are a standard component of care for many patients with CF. Although sampling from clinical OGTTs is typically more sparse than the sampling schedule for this research study, insights into key aspects of altered ISR may lead to improved diagnostic tools to detect early beta cell dysfunction. Since even mild beta cell dysfunction is associated with decreased pulmonary function in patients with CF, early detection of such dysfunction is critical to optimizing CF patient care.

Appendix A. Covariance Functions for Steady-State Conditions. We define $y_0 = y(0)$ with variance

$$(A.1) \quad \text{Var}(y(0)) = \sigma^2 k_\theta(0, 0) = \sigma^2.$$

The covariance function k_θ in (2.5) is chosen to be smooth enough that its mixed second-order partial derivatives coincide, preserving symmetry and allowing for its use in conditioning our ISR inference on an initial derivative value. More precisely, y' is a zero-mean GP whose covariance function is the mixed partial derivative of $\sigma^2 k_\theta$:

$$(A.2) \quad \text{Cov}(y'(t), y'(\tau)) = \sigma^2 \frac{\partial^2 k_\theta(t, \tau)}{\partial t \partial \tau} = \frac{\sigma^2}{\theta^2} \left(1 - \frac{|t - \tau|}{\theta} \right) \exp \left(-\frac{|t - \tau|}{\theta} \right).$$

From (A.2), we can compute the variance

$$(A.3) \quad \text{Var}(y'(0)) = \text{Cov}(y'(0), y'(0)) = \sigma^2 \frac{\partial^2 k_\theta(t, \tau)}{\partial t \partial \tau} \Big|_{(0,0)} = \frac{\sigma^2}{\theta^2}.$$

It remains to specify the covariance between y' and y , which is determined by the appropriate first-order partial derivative of $\sigma^2 k_\theta$, as follows:

$$(A.4) \quad \text{Cov}(y, y') = \frac{\partial k_\theta(t, \tau)}{\partial \tau}.$$

Further, from the symmetry of k_θ ,

$$(A.5) \quad \frac{\partial k_\theta(\tau, t)}{\partial t} = \frac{\partial k_\theta(t, \tau)}{\partial \tau}.$$

Acknowledgments. The authors would like to thank the study participants and their families. The authors would also like to thank Bernard Bialecki for helpful discussions of model discretization.

REFERENCES

- [1] R. M. ABOHTYRA, C. L. CHAN, D. J. ALBERS, AND B. J. GLUCKMAN, *Inferring insulin secretion rate from sparse patient glucose and insulin measures*, *Frontiers in Physiology*, 13 (2022), pp. 1386–1400, <https://doi.org/10.3389/fphys.2022.893862>.
- [2] A. BATTEZZATI, A. MARI, L. ZAZZERON, G. ALICANDRO, L. CLAUT, P. BATTEZZATI, AND C. COLOMBO, *Identification of insulin secretory defects and insulin resistance during oral glucose tolerance test in a cohort of cystic fibrosis patients*, *European Journal of Endocrinology*, 165 (2011), pp. 69–76, <https://doi.org/10.1530/EJE-10-1003>.
- [3] E. BREDI, M. K. CAVAGHAN, G. TOFFOLO, K. S. POLONSKY, AND C. COBELLI, *Oral glucose tolerance test minimal model indexes of beta-cell function and insulin sensitivity*, *Automatica*, 33 (1997), pp. 851–870.
- [4] C. L. CHAN, L. PYLE, T. VIGERS, P. S. ZEITLER, AND K. J. NADEAU, *The relationship between continuous glucose monitoring and ogtt in youth and young adults with cystic fibrosis*, *The Journal of Clinical Endocrinology and Metabolism*, 107 (2021), pp. e548–e560.
- [5] C. COBELLI, C. DALLA MAN, G. TOFFOLO, R. BASU, A. VELLA, AND R. RIZZA, *The oral minimal model method*, *Diabetes*, 64 (2014), pp. 1203–1213.
- [6] G. DE NICOLAO, G. SPARACINO, AND C. COBELLI, *Nonparametric input estimation in physiological systems: Problems, methods, and case studies*, *Automatica*, 33 (1997), pp. 851–870.
- [7] R. P. EATON, R. ALLEN, D. SCHADE, K. ERICKSON, AND J. STANDEFER, *Prehepatic insulin production in man: Kinetic analysis using peripheral connecting peptide behavior*, *The Journal of Clinical Endocrinology and Metabolism*, 51 (1980), pp. 520–528.
- [8] G. FASSHAUER AND M. MCCOURT, *Kernel-Based Approximation Methods using MATLAB*, World Scientific Publishing Co. Pte. Ltd., 2016.

- [9] W. J. HEUETT, B. V. MILLER 3RD, S. B. RACETTE, J. O. HOLLOSZY, C. C. CHOW, AND V. PERIWAL, *Bayesian functional integral method for inferring continuous data from discrete measurements*, Biophysical Journal, 102 (2012), pp. 399–406.
- [10] R. HOVORKA, P. A. SOONS, AND M. A. YOUNG, *ISEC: a program to calculate insulin secretion*, Computational Methods and Programs in Biomedicine, 50 (1996), pp. 253–264.
- [11] L. KJEMS, A. VØLUND, AND S. MADSBAD, *Quantification of beta-cell function during ivgtt in type ii and non-diabetic subjects: assessment of insulin secretion by mathematical methods*, Diabetologia, 44 (2001), pp. 1339–1348.
- [12] C. LEWIS, S. M. BLACKMAN, A. NELSON, E. OBERDORDER, D. WELLS, J. DUNITZ, W. THOMAS, AND A. MORAN, *Diabetes-related mortality in adults with cystic fibrosis. role of genotype and sex*, American Journal of Physiology: Endocrinology and Metabolism, 298 (2010), pp. E440–E448.
- [13] W. LIU, C. HSIN, AND F. TANG, *A molecular mathematical model of glucose mobilization and uptake*, Mathematical Biosciences, 221 (2009), pp. 121–129.
- [14] A. MARI AND E. FERRANNINI, *β -cell function assessment from modelling of oral tests: an effective approach*, Diabetes, Obesity and Metabolism, 10 (2008), pp. 77–87, <https://doi.org/10.1111/j.1463-1326.2008.00946.x>.
- [15] A. MARI, A. TURA, A. GASTALDELLI, AND E. FERRANNINI, *Assessing insulin secretion by modeling in multiple-meal tests: role of potentiation*, Diabetes, 51 (2002), pp. S221–S226, <https://doi.org/10.2337/diabetes.51.2007.s221>.
- [16] A. MORAN, C. BRUNZELL, R. C. COHEN, M. KATZ, B. C. MARSHALL, G. ONADY, K. A. ROBINSON, K. A. SABADOSA, A. STECENKO, B. SLOVIS, AND THE CFRD GUIDELINES COMMITTEE, *Clinical Care Guidelines for Cystic Fibrosis-Related Diabetes: A position statement of the American Diabetes Association and a clinical practice guideline of the Cystic Fibrosis Foundation, endorsed by the Pediatric Endocrine Society*, Diabetes Care, 33 (2010), pp. 2697–2708.
- [17] A. MORAN, J. DUNITZ, B. NATHAN, A. SAEED, B. HOLME, AND W. THOMAS, *Cystic fibrosis-related diabetes: current trends in prevalence, incidence, and mortality*, Diabetes Care, 32 (2009), pp. 1626–1631, <https://doi.org/10.2337/dc09-0586>.
- [18] G. PILLONETTO, G. SPARACINO, AND C. COBELLI, *Numerical non-identifiability regions of the minimal model of glucose kinetics: superiority of bayesian estimation*, Mathematical Biosciences, 184 (2003), pp. 53–67.
- [19] C. PIONA, S. VOLPI, C. ZUSI, E. MOZZILLO, A. TOSCO, A. FRANZESE, V. RAIA, M. L. BOSELLI, M. TROMBETTA, M. CIPOLLI, R. C. BONADONNA, AND C. MAFFEIS, *Glucose tolerance stages in cystic fibrosis are identified by a unique pattern of defects of beta-cell function*, The Journal of Clinical Endocrinology and Metabolism, 106 (2021), pp. 1793–1802, <https://doi.org/10.1210/clinem/dgaa932>.
- [20] K. S. POLONSKY, B. GIVEN, AND E. VAN CAUTER, *Twenty-four-hour profiles and pulsatile patterns of insulin secretion in normal and obese subjects*, Journal of Clinical Investigation, 81 (1988), pp. 442–448.
- [21] M. SCHIAVON, D. HERZIG, M. HEPPRICH, M. Y. DONATH, L. BALLY, AND C. DALLA MAN, *Model-based assessment of c-peptide secretion and kinetics in post gastric bypass individuals experiencing postprandial hyperinsulinemic hypoglycemia*, Frontiers in Endocrinology, 12 (2021), <https://doi.org/10.3389/fendo.2021.611253>, <https://www.frontiersin.org/articles/10.3389/fendo.2021.611253>.
- [22] S. SHEIKH, L. GUDIPATY, D. D. DE LEON, D. HADJILIADIS, C. KUBRAK, N. K. ROSENFELD, S. C. NYIRJESY, A. J. PELECKIS, S. MALIK, D. STEFANOVSKI, M. CUCHEL, R. C. RUBENSTEIN, A. KELLY, AND M. R. RICKELS, *Reduced β -cell secretory capacity in pancreatic-insufficient, but not pancreatic-sufficient, cystic fibrosis despite normal glucose tolerance*, Diabetes, 66 (2017), pp. 134–144.
- [23] M. L. STEIN, *Interpolation of spatial data: some theory for kriging*, Springer Science & Business Media, 1999.
- [24] I. M. TOLIĆ, E. MOSEKILDE, AND J. STURIS, *Modeling the insulin-glucose feedback system: The significance of pulsatile insulin secretion*, Journal of Theoretical Biology, 207 (2000), pp. 361–375.
- [25] K. L. TOMMERDAHL, J. T. BRINTON, T. VIGERS, M. CREE-GREEN, P. S. ZEITLER, K. J. NADEAU, AND C. L. CHAN, *Delayed glucose peak and elevated 1-hour glucose on the oral glucose tolerance test identify youth with cystic fibrosis with lower oral disposition index*, Journal of Cystic Fibrosis, 20 (2021), pp. 339–345.
- [26] E. VAN CAUTER, F. MESTREZ, J. STURIS, AND K. S. POLONSKY, *Estimation of insulin secretion rates from c-peptide levels. comparison of individual and standard kinetic parameters for c-peptide clearance*, Diabetes, 41 (1992), pp. 368–377.