

<https://doi.org/10.1038/s41540-025-00570-6>

Conditional universal differential equations capture population dynamics and interindividual variation in c-peptide production

Max de Rooij^{1,2}✉, Natal A. W. van Riel^{1,2} & Shauna D. O'Donovan^{1,2}

Universal differential equations (UDEs) are an emerging approach in biomedical systems biology, integrating physiology-driven mathematical models with machine learning for data-driven model discovery in areas where knowledge of the underlying physiology is limited. However, current approaches to training UDEs do not directly accommodate heterogeneity in the underlying data. As a data-driven approach, UDEs are also vulnerable to overfitting and consequently cannot sufficiently generalize to heterogeneous populations. We propose a conditional UDE (cUDE) where we assume that the structure and weights of the embedded neural network are common across individuals, and introduce a conditioning parameter that is allowed to vary between individuals. In this way, the cUDE architecture can accommodate inter-individual variation in data while learning a generalizable network representation. We demonstrate the effectiveness of the cUDE as an extension of the UDE framework by training a cUDE model of c-peptide production. We show that our cUDE model can accurately describe postprandial c-peptide levels in individuals with normal glucose tolerance, impaired glucose tolerance, and type 2 diabetes mellitus. Furthermore, we show that the conditional parameter captures relevant inter-individual variation. Subsequently, we use symbolic regression to derive a generalizable analytical expression for c-peptide production.

Many current medical treatments and interventions have been developed and tested in clinical trials involving cohorts of individuals. Whereas inter-individual variability in subjects included in clinical trials is typically strongly characterized, prescription of treatment often assumes that patients receiving respond similarly to the average response in these clinical trials¹. However, this assumption inherently neglects physiological and environmental differences between people, such as genetic variants or acquired exposures that may mediate disease risk or response to treatment. Furthermore, the observed inter-individual variation may not be due to just natural variability, but this variation may be indicative of disease progression. Specifically, in the development of type 2 diabetes mellitus (T2DM), the progressive decline of β -cell function, responsible for insulin release in response to glucose, is a key characteristic of disease development². Furthermore, residual β -cell function is indicative of treatment response and can therefore aid in treatment selection^{3,4}.

In recent years, personalized medicine, where treatments are tailored based on specific characteristics, such as genetics⁵ or body composition⁶, has emerged as a promising approach to improve health outcomes. In particular, in the field of oncology, machine learning is increasingly being used to map large datasets to clinical outputs, to identify more optimal treatment strategies based on genetic data⁷ or machine learning-assisted analysis of tumor biopsies⁸. Research suggests that these more personalized treatment regimens have the potential to improve the long-term prognosis of patients⁹. However, the direct application of machine learning methods that have shown success in precision oncology to other medical disciplines has been hampered by smaller sample sizes and a lack of publicly available clinical trial data^{10,11}.

The advantages of these purely data-driven approaches that have been used successfully in precision oncology are that they allow flexible incorporation of various types and sources of data for accurate model output. However, a downside of this flexibility in machine learning models is that the volume of data required to train machine learning models is relatively

¹Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands. ²Eindhoven Artificial Intelligence Systems Institute (EAISI), Eindhoven University of Technology, Eindhoven, The Netherlands. ✉e-mail: m.d.rooij@tue.nl

large¹². The comparatively small sample sizes collected in human clinical trials have greatly hampered the widespread deployment of machine learning to biomedical problems¹⁰. In addition, these machine learning models can lack interpretability, particularly in the case of larger neural networks¹³. This interpretability can be retained by using inherently predictable models, such as ordinary logistic regression, for structured data with meaningful features.

Alternatively, in cases where individual data is limited but biological knowledge is abundant, systems of differential equations are constructed to describe biological processes. These physiologically-based mathematical models (PBMMs) are powerful tools to disentangle the complexity of the physiological basis underlying medical measurements^{14,15}. Previous research has demonstrated that the estimation of model parameters in PBMMs from individual measurements can yield an accurate and interpretable explanation of the inter-individual biological variation¹⁶. While PBMMs are beneficial for studying biological systems, building and validating accurate PBMMs requires a profound understanding of the underlying physiology and can be time-consuming. Consequently, these models typically have a limited scope^{15,17}. Additionally, as PBMMs are constructed manually, unwanted bias can be introduced, especially when complex nonlinear biological behaviours may be approximated with comparatively simple terms.

A promising emerging area of research focuses on the combination of highly plastic machine learning approaches with physiological knowledge in the form of mechanistic models to produce a hybrid model that can be trained with fewer learning examples. In recent years, multiple hybrid frameworks have been proposed. Physics-informed neural networks (PINNs)¹⁸ are an example, where the loss function of a neural network is supplemented with a set of equations to ensure that the neural network not only fits the data well, but also adheres to known physical laws. In this work, we use the Universal Differential Equations (UDE) framework¹⁹, where the known components of a biological system are described by parameterized differential equations and a neural network is incorporated into the equations to account for the unknown components. These UDE models have been shown to be applicable to various biological systems, including the deployment to infer the glucose appearance of a meal²⁰, as well as a STAT5 dimerization model²¹. Furthermore, the resulting trained neural networks can be reduced to analytical expressions using a technique called symbolic regression¹⁹.

The application of UDE models in biomedical contexts such as learning the average rate of glucose appearance from a meal has been explored. However, current conventional training of UDEs cannot directly accommodate inter-individual variability, which is ubiquitous in biomedical data. Although it is possible to train the model on the average data of a population, as has been done in the past, such models are not expected to generalize well to the individual data. Alternatively, it is possible to train a model for each individual separately. However, this approach has some drawbacks. First, often only limited measurements are available for each individual, making estimation of neural network parameters on individuals highly sensitive to measurement noise, increasing the risk of overfitting. Furthermore, the black-box nature of neural networks complicates the comparison of trained neural networks between individuals.

In this work, we propose an extension of the UDE framework, termed conditional UDEs (cUDEs), where trainable person-specific parameters are added as input to the neural network to account for between-subject variability, and the weights of the neural network are assumed to be common across the entire population. In this way, variability between subjects is forced into these conditional input parameters, while the neural network parameters learn the global behaviour of the system.

Here, we applied a cUDE model to characterize the insulin production capacity of pancreatic β -cells in individuals with normal glucose tolerance, impaired glucose tolerance and T2DM. Our results demonstrate that the conditional universal differential framework derives an accurate representation of the inter-individual variation in c-peptide production. Furthermore, we show that this subject-specific conditioning parameter is strongly correlated with the gold standard hyperglycemic clamp measure of insulin production capacity. We then derived an analytical expression from

the conditionally trained network using symbolic regression and showed that the learned function not only described c-peptide production for people with normal glucose tolerance, impaired glucose tolerance, and T2DM, but also generalized to describe individual c-peptide production in an independent human trial.

Results

Conventionally trained UDE does not generalize across population

To investigate the ability of a conventionally trained UDE to generalize to meal responses in a population of individuals, a universal differential equation of c-peptide production and kinetics was initially trained on the average meal response. The UDE model is based on a two-compartment ordinary differential equation model describing c-peptide kinetics in the plasma and interstitial space by van Cauter et al.²². Here, the van Cauter model was extended by introducing a fully connected neural network to represent c-peptide production in the pancreas (Fig. 1a).

The change in plasma glucose concentration relative to the fasting value at time is provided as input to the neural network t , defined as

$$G(t) = G^{\text{pl}}(t) - G^{\text{pl}}(0)$$

where $G^{\text{pl}}(t)$ is the plasma glucose value at time t . The output of the neural network is the rate of c-peptide production $P(t)$.

To train the model, demographic data and plasma glucose and c-peptide trajectories were used from 117 people from a study by Okuno et al.²³, labeled the Ohashi dataset, as the data was retrieved from a paper by Ohashi et al.²⁴. The data set encompassed three distinct subgroups: people with normal glucose tolerance (NGT), impaired glucose tolerance (IGT), and type 2 diabetes mellitus (T2DM) (Supplementary Fig. 1). For estimating neural network weights and biases, average c-peptide measurements were used from a training set containing 70% of the individuals. The weights and biases from the neural network trained on the average response were then used in combination with the glucose values and kinetic parameters to predict postprandial c-peptide values. The simulation errors for the individuals in the train and test sets are shown in Fig. 1b, showing comparable performance for the normal glucose tolerance and impaired glucose tolerance groups, but a strong reduction in performance in the T2DM group.

Figure 1 c–e shows the resulting UDE fits, using the average data from each glucose tolerance condition as input. From this figure, we can observe that the UDE model generally fits the mean data within one standard deviation, with the exception of the final two time points in the IGT group. However, the model underestimates c-peptide production in the NGT and IGT groups, while overestimating c-peptide production in the T2DM group. This indicates the inability of the single universal differential equation trained on the average response data to account for the progressive decline in β -cell function observed in the progression from NGT towards T2DM.

Conditional universal differential equation model of c-peptide kinetics

To capture the inter-individual variability, an additional input parameter was added to the neural network, resulting in a conditional UDE model (cUDE). Consequently, the neural network in this cUDE model has two inputs. The first input of the cUDE network is the relative plasma glucose concentration at time t , as in the conventional UDE. The second input is a trainable parameter β_i that accounts for the variability between individuals in the production of c-peptide. The output of the neural network ($P(t)$) is the c-peptide production at time t (Fig. 2a).

Figure 2b depicts the process of training the conditional UDE. Model selection and training are performed on a subset of 70% the dataset, labeled the 'train set'. The weights and biases of the neural network for the whole population are trained together with the individual parameters of the train set, obtaining 25 candidate models from 25 initializations of the optimization. A validation set is used, where only the individual parameters are estimated, to select the best-performing model from these 25 candidate

Fig. 1 | Modelling c-peptide production with a conventionally trained universal differential equation. **a** Schematic overview of the van Cauter model of c-peptide kinetics, depicting the location of the neural network that describes the production of c-peptide ($P(t)$) depending on plasma glucose (G_{pl}). The blue circles indicate the c-peptide state variables (C_{pl} and C_{int} for the plasma and interstitial fluid compartments respectively). The green circle depicts the plasma glucose level. Solid arrows represent fluxes, and dashed arrows indicate stimulation. Each flux arrow is labeled with their respective kinetic parameter. **b** Mean squared error (MSE) distributions of the UDE model, trained on average response data, on each individual in the used dataset, split by train and test set, and grouped by glucose tolerance status. **c–e** Mean (circles) and standard deviations (error bars) of the data, and UDE model predictions (solid lines) given the mean data per glucose tolerance condition.

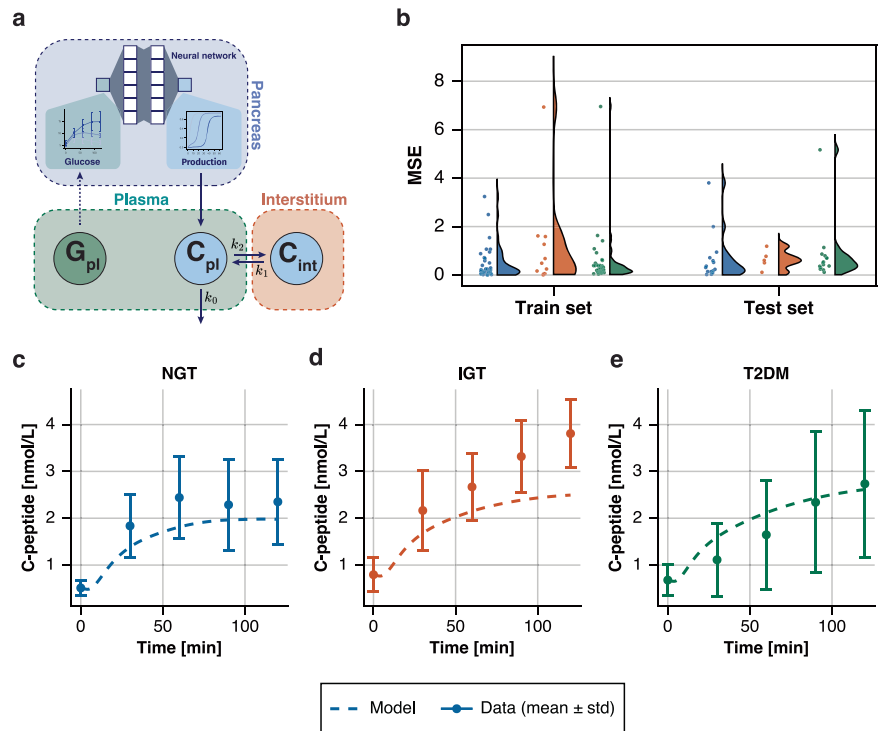
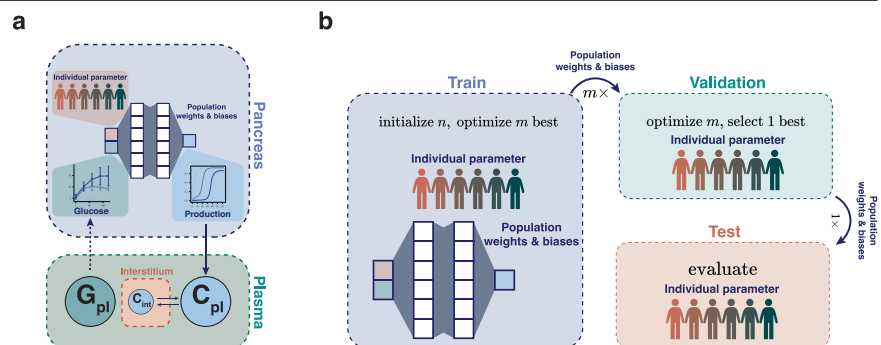


Fig. 2 | Structure and training procedure of the conditional universal differential equation model used to infer postprandial c-peptide. **a** Schematic of the conditional UDE model, and the neural network used to estimate personalized c-peptide production. The time-dependent plasma glucose value and a person-specific parameter controlling for inter-individual variability in dose response are inputs to the neural network. The weights and biases of the neural network are estimated population-wide. **b** Illustration of the training procedure. The dataset is split into a train (49%), validation (21%), and test set (30%). In the train set, both population and individual parameters are estimated. In the validation and test sets, population parameters in the neural network are fixed to the trained values and only the person-specific conditioning parameters are estimated from data.



models. The model is then evaluated on a separate test set, where, in the same way as in the validation set, the individual parameters are estimated, while the neural network parameters are kept constant.

cUDE derives generalizable c-peptide production across population

Figure 3a–c visualizes the cUDE simulation of plasma c-peptide for the individuals in the test set with the median error value for each glucose tolerance condition, showing a good concordance with the measured c-peptide data. This figure demonstrates that the same neural network weights and biases, in combination with a subject-specific conditional parameter, can simulate glucose-driven c-peptide production while accounting for a large part of the inter-individual variability of the c-peptide production. The confidence regions for the model simulations are computed from the likelihood profiles, shown in Supplementary Fig. 6. All data points for the individuals are contained within these confidence regions, with the exception of the final time point in the NGT case. Furthermore, the confidence regions for the NGT and IGT individuals are both larger than the confidence region in

the T2DM individual. All test fits are shown in Supplementary Fig. 4. The empirical distributions of the conditional parameters were computed for each glucose tolerance condition, and included in Supplementary Fig. 5a, where Supplementary Fig. 5b–d contain simulated c-peptide curves for each glucose tolerance condition, showing that these curves closely match with the c-peptide data of each glucose tolerance condition.

Furthermore, the distribution of model error values across the three glucose tolerance groups is shown in 3 d. Compared to the conventional UDE (Fig. 1b), the distributions are narrower, especially for the T2DM group. The resulting model fits and error distributions are comparable to the model fits and errors in the train set, which can be found in Supplementary Fig. 3. In addition, training the cUDE model on various fractions of the train set showed that a train set size of around 29 individuals is already sufficient for training a model, with a comparable mean test error to the current cUDE model (Supplementary Fig. 7).

In supplementary section Supplementary Note 2, we demonstrate that the ability of the cUDE architecture to learn a generalizable model from data with large systematic heterogeneity by applying the approach to a second

Fig. 3 | Model fits of the conditional UDE (cUDE) model on the test data. **a–c** Model fit of the individuals with median error value within each glucose tolerance group. Visualization of all model fits for all individuals in the test set can be found in Supplementary Fig. 4. Circles indicate the measured c-peptide levels from each individual and solid lines represent the model fits. Dotted lines represent the 95% confidence intervals on the model fits based on the likelihood profiles, defined according to⁴⁵. **d** Distribution of mean squared error values for model fits for all subjects in the test subset, separated by glucose tolerance status group.

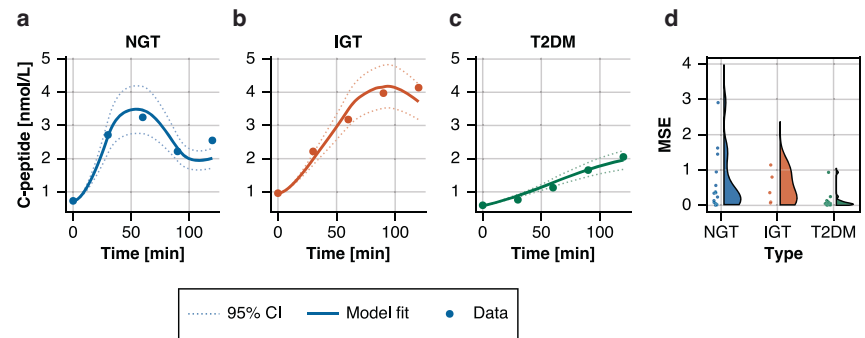
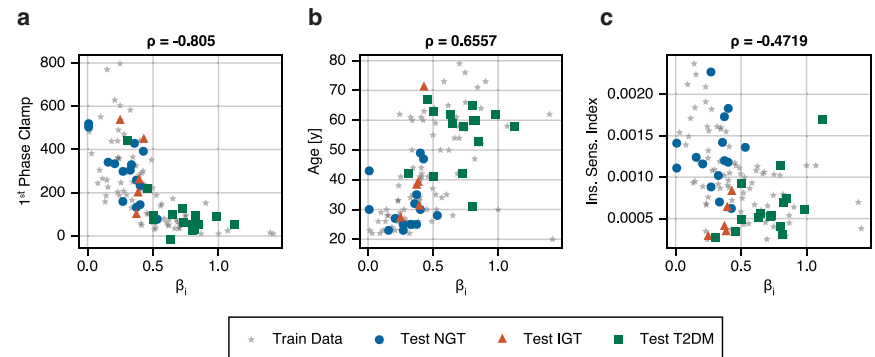


Fig. 4 | Spearman correlation of conditional parameter β_i with independent phenotypic measurements for the individuals. **a** Spearman correlation of the conditional parameter with the first-phase insulin production during a hyperglycemic clamp (Supplementary Fig. 2). **b** Correlation of the conditional parameter with age in years. **c** Correlation of the conditional parameter with the insulin sensitivity index measured from a hyperinsulinemic-euglycemic clamp test.



example system simulated using a different mathematical model. As with the c-peptide model, the conditional parameter showed strong correlation with the prescribed parameter values used to simulate the data. (Supplementary Note 2).

Conditional training parameter captures inter-individual variation

To investigate the interpretability of the conditional parameter, personalized conditional parameters were compared with subject characteristics, including BMI, age, body weight, and clamp-based measurements of insulin sensitivity and insulin production capacity.

In Fig. 4, the strongest Spearman correlation of -0.805 is observed with the first phase of insulin production measured using the hyperglycemic clamp, the gold standard measure of insulin production (Fig. 4a). A moderate correlation is seen with age (b), while the insulin sensitivity index, measured using a hyperinsulinemic-euglycemic clamp (c), has the lowest correlation of the three.

The correlations with body weight and body mass index are low, while the correlations with other measures of insulin production are high, which is shown in the supplementary figure Supplementary Fig. 10.

Symbolic regression derives a generalizable analytical expression of c-peptide production

As the neural network model remains a black box model, we also sought to replace the neural network with a more interpretable analytical expression. The symbolic regression approach proposed by Cranmer et al.²⁵ was applied to data sampled from the trained neural network. Subsequently, the derived analytical expression was simplified manually, reducing several fixed constants to a single term (see Supplementary Note 1 for a detailed derivation). The resulting expression resembles Michaelis-Menten kinetics and is given as

$$P(G^{pl}(t)|k_M) = \begin{cases} 1.78 \cdot \frac{G^{pl}(t) - G^{pl}(0)}{k_M + G^{pl}(t) - G^{pl}(0)} & \text{if } G^{pl}(t) \geq G^{pl}(0) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Here, k_M is a trainable parameter, qualitatively equivalent to the β_i parameter learned by the cUDE. (see Supplementary Note 1 for more details on the numerical relation between β_i and k_M). The dose-response curves for the neural network and the learned expression are depicted in supplementary figure Supplementary Fig. 12.

To evaluate the performance of this learned analytical expression for c-peptide production the neural network of the cUDE was replaced with equation (1). The fully analytical model was then fit to the measured c-peptide data for all individuals by estimating a value for k_M .

Figure 5 a–c visualizes the model fits of the analytical model model for the individuals corresponding to the median error values per glucose tolerance group. As seen with the cUDE model, the model derived via symbolic regression agrees well with the data across all three groups. The distribution of model fit errors per group (Fig. 5d) also shows comparable distributions to the model fit errors obtained for the cUDE model. Furthermore, the correlations of the estimated k_M value with insulin production, age, and insulin sensitivity, as shown in Fig. 5e–g, are similar to the previous results obtained with the cUDE model and again display a high correlation with insulin production as measured with the hyperglycaemic clamp. Profile likelihood analysis was performed on the parameter k_M for each individual to test whether it was identifiable from the data. (Supplementary Fig. 6)

Finally, to demonstrate generalizability of the model derived from symbolic regression, the analytical model was fitted to glucose and c-peptide measurements collected during an OGTT from a previously unseen dataset.

The model fits for the individuals at the 25th, 50th and 75th percentiles of the mean squared error are shown in Fig. 6a–c respectively. In all three models, the curve shows high concordance with the data. Moreover, despite the original cUDE model being trained for data up to 120 min, the learned analytical term can also reliably simulate plasma concentrations of c-peptide up to 240 min postprandially. The distribution of model errors is shown in Fig. 6d, indicating that high-quality model fits could be obtained for a large part of the twenty individuals.

Fig. 5 | Fit of the analytical model derived using symbolic regression to measured data. **a–c** Model fit for individuals with the median error value for each glucose tolerance condition. Model fits are shown with the solid line, measured c-peptide values are indicated with the circles. The model simulations for the 95% confidence intervals on the parameters are shown in dashed lines. **d** Model error value distributions split by glucose tolerance condition. **e–g** Correlation of personalized k_M estimate with **e** the first-phase insulin production measured during the hyperglycaemic clamp, **f** age (years), and **g** the insulin sensitivity index measured using a hyperinsulinemic-euglycemic clamp test.

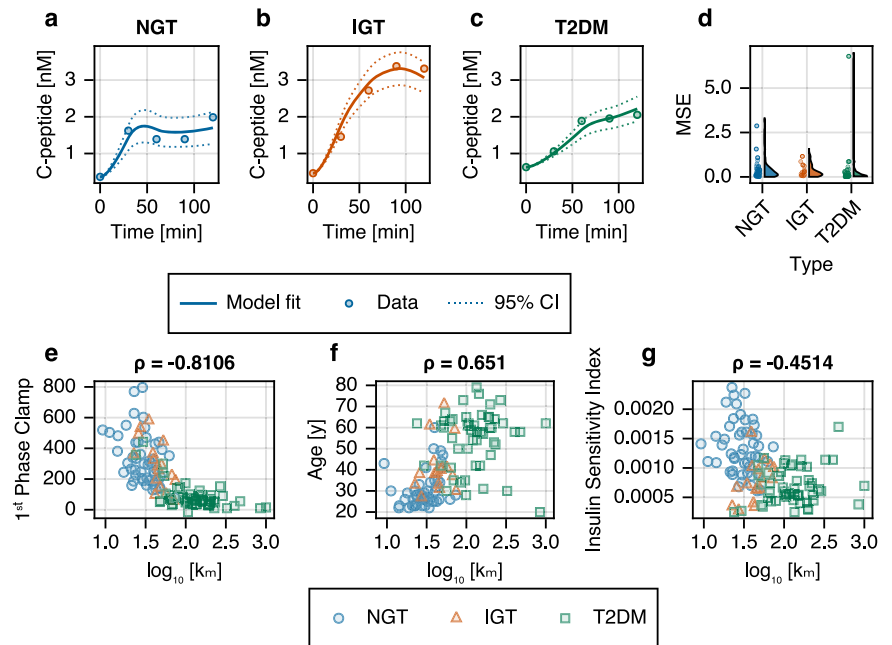
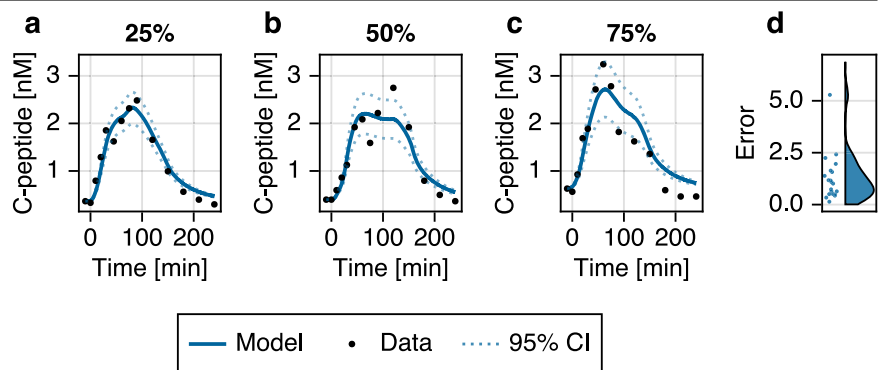


Fig. 6 | Model fits and errors for the c-peptide model derived using symbolic regression on the Fujita dataset⁴³. **a–c** Model fits for the individuals at the 25th, 50th and 75th percentiles of the model error distribution respectively. Measured c-peptide values are indicated with the black circles, the simulated model fits are shown with the solid lines. Dashed lines indicate simulations with the 95% confidence intervals on the parameters. **d** Raincloud plot of the model errors for the entire Fujita dataset.



Discussion

In this work, we introduced conditional universal differential equations (cUDEs) as an extension of the universal differential equation framework that facilitates simultaneous data-driven model discovery and model personalization. We then applied this technique to uncover a novel index of inter-individual variation in c-peptide, and by extension, insulin production in a human population with diverse glucose tolerance status.

Our results show that cUDE models can accurately estimate a missing c-peptide production term from the data. More importantly, by accommodating the large inter-individual variation in plasma glucose and c-peptide level, the cUDE learns a model that can be generalized across individuals with different glucose tolerance status. In contrast, the classical UDE was unable to capture difference in beta-cell capacity that are indicative of glucose tolerance status. Investigating individual model fits, the trained cUDE was unable to describe the c-peptide measurements of a single individual (Supplementary Fig. 4 individual 10) from the test set. However, this individual showed a strong discordance between the measured glucose and c-peptide data with measured plasma glucose only increasing 60 minutes after ingestion of the glucose solution. This unexpected plasma glucose response may potentially be explained by the effect of incretin hormones such as GLP-1 or GIP. These incretin hormones are produced in response to an increase in glucose level in the intestine and activate insulin and c-peptide production^{26,27}. In this study, these hormones are not measured and are a

potential additional source of inter-individual variability in c-peptide production. Should time series of incretin hormones become available in the future, the cUDE framework could be reapplied without strong modifications to further learn the role of these incretin hormones in c-peptide production. However, in the current model, where only glucose is provided as the stimulus for c-peptide release, the majority of model fits showed a strong agreement with plasma measurements, suggesting that glucose is the primary driver of c-peptide production²⁸.

Furthermore, by constraining the weights and biases of the neural network to be the same for the entire population, the free conditional parameters capture the inter-individual variation which enables the direct comparison between individuals. By comparing the conditional parameters resulting from the c-peptide model with a range of independent measures of metabolic health, we have shown that the conditional parameter strongly correlates with metrics of insulin secretion measured using the hyperglycemic clamp method, the current gold standard measure of insulin production capacity. Furthermore, the lack of a strong correlation with the insulin sensitivity index indicates that the conditional parameter specifically targets the c-peptide and insulin production capacity, and not just a general deterioration in metabolic resilience. The moderate correlation observed with age may have two causes. Firstly, the conditional parameter has been shown to describe the progressive decline in β -cell function, with higher values in people with T2DM. The age distribution was different between

each glucose tolerance condition, with the ages of the T2DM group being significantly different than the NGT individuals (Mann-Whitney U test, $p < 10^{-10}$). Secondly, part of this correlation may also originate from the known natural decline of β -cell function with aging²⁹. We have also trained a cUDE model including age as an additional input to the model to investigate the effect of correcting for age on the correlation with the first-phase clamp indices and the curve-fitting performance. However, the correlation of the conditional parameter with the clamp index reduced slightly, and no notable improvement was observed regarding curve fitting performance. (Supplementary Fig. 11) We suspect that the bias of the dataset concerning the age of individuals in each subgroup may influence the ability of the neural network to correctly estimate the true age effect on insulin production capacity, and we would require a larger dataset with a better representation of this natural age effect for an improved separation of the age effect and the diabetes progression effect. While the inclusion of age as an additional covariate did not improve the results in the model used in this work, this feature of cUDEs could also be used in different applications, for example to introduce relevant phenotypic characteristics, such as sex, smoking status, or family history of disease into the UDE models. This approach would produce a hybrid model that can integrate these features into a mechanistic model to improve the prediction of disease risk or treatment response, as proposed in the case of ventricular tachycardia in ref. 30.

In order to learn a generalizable model of c-peptide production, a sufficiently heterogeneous dataset is required. Here, we used data from the Ohashi data set consisting of individuals with normal glucose tolerance, impaired glucose tolerance and T2DM. However, it is not essential to have very large data sets. In Supplementary Fig. 7 we show that using data from 29 individuals did not strongly increase the test error of the cUDE model, provided that the proportion of NGT, IGT and T2DM was maintained. Although the amount of data required also depends on the complexity of the model to be learned, and thereby the neural network size and the amount of inputs and outputs, the cUDE is relatively data-efficient, compared to fully data-driven methods, which typically require thousands of samples^{31,32}. Additionally, we have shown the applicability of the cUDE model in a simulated example with just 37 individuals, showing that the conditional parameter strongly associates with the variability introduced in the simulated data. (Supplementary Note 2).

Furthermore, we show that the interpretability of the cUDE model can also be further increased through the use of symbolic regression. For symbolic regression, we have used a genetic algorithm, which is non-deterministic and may produce variable results upon repeated runs. This can be mitigated by letting the algorithm run through sufficient iterations, which will eventually lead to model convergence. However, this required number of iterations (25,000 in this work) is problem dependent, and for larger problems, more iterations are required, which should be taken into account when applying symbolic regression based on genetic algorithms. In this work, the use of a limited number of allowed operators based on knowledge of previously built ODE models in systems biology greatly reduced the search space, allowing for the discovery of an interpretable model. However, some detail in the dose-response relationship is lost when comparing the analytic expression to the neural network (Supplementary Fig. 12). Despite the loss of some of the details in the dose response, the derived analytical equation demonstrates generalizability beyond the original dataset, as shown by fitting the derived analytical model to normoglycemic individuals from a previously unseen dataset. Furthermore, we demonstrate that the derived model originally trained on 120 minutes of data can successfully simulate model behaviour over 240 minutes.

Several models of insulin and c-peptide production in response to glucose have been proposed in the literature. From models such as Maas et al.³³ that use a complex PID controller or the detailed model of exocytosis used by Ha et al.³⁴ to simple linear mass-action kinetics presented in Hovorka et al.³⁵. The Michaelis-Menten term for c-peptide production derived from data in this study is similar to the insulin production term used in the model by Topp et al.³⁶, which uses a Hill function with a Hill

coefficient of 2. This acts as a form of validation that the model we derived from symbolic regression is a physiologically plausible model of c-peptide production.

A limitation of this study is that both the Ohashi and Fujita datasets contain only people of Japanese descent. Although previous work has provided evidence for similar β -cell responsiveness across all glucose tolerance states³⁷, it is necessary to further validate the trained model on more diverse populations. In addition, the derived model has only been tested on OGTT responses. Especially considering the effect of amino acids on insulin and c-peptide production³⁸, the model may not be able to accurately describe responses to more complex meals. However, despite these limitations in the learned c-peptide model, we demonstrate that the cUDE approach outperforms current UDE approaches in learning a generalizable model that incorporates biologically relevant inter-individual variation.

While we demonstrated that the conditional parameters were identifiable in almost all individuals (109 of 117 individuals, Supplementary Fig. 6), we also demonstrated that the conditional parameters are only identifiable when the neural network parameters are fixed, as can be seen from Supplementary Fig. 8. In this figure, we visualized β against the first phase of the hyperglycemic clamp for all models resulting from the various initializations of the neural network parameters during training. When comparing the conditional parameters trained in multiple initializations of the cUDE training we see that the neural network learns either a positive or negative association with the first-phase clamp index. However, a consistently strong association is derived across models. Furthermore, while a linear relationship between the conditional parameters of two models is not guaranteed, due to the nonlinearity of the neural network, comparing the parameters of two models does results in a high correlation. This high correlation, in combination with narrow spread of points suggests an algebraic relationship (Supplementary Fig. 9). This effect, however, does pose a challenge concerning the use of ensemble UDE models for increased robustness³⁹. This challenge can potentially be remedied using dimensionality reduction techniques, such as principal component analysis, to align common patterns within conditional parameters, but this requires further investigation.

Furthermore, if multiple conditional parameters were to be used, the nonlinearity of the neural network can cause these to become correlated, and mutually unidentifiable. Possibly, in case of multiple conditional parameters, regularization could be applied to penalize correlations between the conditional parameters to ensure orthogonality. However, assessment of identifiability is still only possible after fixing the neural network weights and biases.

In our current training regimen, we train the biases and weights of the neural network using the whole training data set, while the conditional parameters are trained independently for each individual using a maximum likelihood approach. Nonlinear mixed effects (NLME) modelling is an alternative approach to model parameterization that simultaneously accounts for both inter- and intra-individual variability⁴⁰. By representing inter-individual variability through random effects, NLME models enable scalable estimation via the population likelihood, integrating out individual-level parameters. Recent advances, such as neural network-based NLME extensions, incorporate random effects as neural inputs to capture population heterogeneity, typically under the assumption of normally distributed effects⁴¹. We show that NLME based training of the cUDE model is equally possible and yields similar results to the original approach taken in this work. While the correlation with hyperglycemic clamp remains strong when estimating the parameters using a NLME structure, the accuracy of the model fit is reduced in some individuals of the T2DM group, as their parameters regress towards the population mean. (Supplementary Note 3) Due to this reduced accuracy, we used a traditional frequentist approach for estimating parameters, but in some cases, depending on the research question being addressed, NLME estimation combined with the cUDE model structure may provide more useful results.

In conclusion, we present cUDEs as an effective extension to the UDE framework that can be used to learn a generalizable representation of

missing dynamics from a heterogeneous dataset. The cUDE works under the main assumption that the dynamic system underlying the data is common to all samples, while only a limited set of parameters is necessary to capture the differences between samples. This setup makes the cUDE especially suited to biological challenges, where inter-individual variability is both ubiquitous and often physiologically relevant. Here, we show that the conditional parameter in the cUDE model for c-peptide is interpretable as a physiologically relevant index, capturing the inter-individual variability in c-peptide production as validated by comparison with the hyperglycemic clamp. Although this study demonstrates the application of the cUDE model in a specific application, the cUDE framework is also usable in several other medical disciplines where mathematical models are abundant, such as cardiovascular medicine, neurology, and infectious diseases. The ability of the cUDE model to learn a model that can generalize, capture relevant and interpretable inter-individual variation, and to be trained with limited number of learning examples are key features that demonstrate its potential to support model- and data-driven precision healthcare.

Methods

Ohashi dataset

The Ohashi dataset was obtained from Ohashi et al.^{24,42}, and originally collected by Okuno et al.²³. The original study was approved by the ethics committee of the Kobe University Graduate School of Medicine and was registered with the University Hospital Medical Information Network (UMIN000002359). Written informed consent was obtained for all subjects.

As described in⁴², 50 subjects with normal glucose tolerance (NGT), 18 subjects with impaired glucose tolerance (IGT), and 53 subjects with type 2 diabetes (T2DM) participated in the study. The characteristics of the subjects for each group are shown in Table 1.

All subjects underwent a 75-gram oral glucose tolerance test, as well as a consecutive hyperglycemic and hyperinsulinemic-euglycemic clamp test. Both tests were performed on separate mornings after an overnight fast.

In the 75g-OGTT, following an overnight fast, blood samples were collected before and 30, 60, 90 and 120 minutes after ingestion of the glucose solution. Plasma glucose and serum insulin and c-peptide concentrations were measured in each sample.

Hyperglycemic clamp and hyperinsulinemic-euglycemic clamp tests were performed consecutively. The hyperglycemic clamp began with an intravenous infusion of a glucose bolus of 9622 mgm⁻² within 15 minutes, followed by a variable dose of glucose to keep plasma glucose levels at 200 mg dL⁻¹ for 90 minutes. Blood samples were collected before and at 5, 10, 15, 60, 75, and 90 minutes after glucose infusion. In each blood sample, plasma glucose and serum insulin and c-peptide were measured. The hyperinsulinemic-euglycemic clamp test was then performed by intravenous infusion of regular human insulin at 1.46 mUkg⁻¹ min⁻¹ to obtain a serum insulin concentration of 600 pmolL⁻¹. Plasma glucose concentration was kept at 90 mg dL⁻¹ by variable glucose infusion for 120 minutes²³.

Insulin secretion indices were defined as the incremental area under the insulin concentration curve during the hyperglycemic clamp:

$$S(T_1, T_2) = \int_{T_1}^{T_2} (I(t) - I(0))dt \quad (2)$$

Where the insulin secretion during the first-phase is defined as $S(0, 10)$, during the second phase as $S(10, 90)$ and the total insulin secretion as $S(0, 90)$. The insulin sensitivity index (ISI) is calculated from the hyperinsulinemic-euglycemic clamp by dividing the mean measured glucose infusion rate during the last 30 minutes of the test by the product of plasma glucose and serum insulin levels at the end of the clamp ($t = 120$).

Fujita dataset

The Fujita dataset was obtained from Fujita et al.⁴³. Written informed consent was obtained for all subjects.

As described in⁴³, 20 subjects with normal glucose tolerance (NGT) participated in the study. Subject characteristics for each group are shown in

Table 1 | Subject characteristics from the Ohashi dataset^{23,24,42} after exclusion of subjects with missing data, and the Fujita dataset⁴³

	Ohashi			Fujita
	NGT	IGT	T2DM	—
number of individuals	49	17	51	20
male/female	22/27	10/7	31/20	14/6
age (y) ^a	30 ± 8.5	41 ± 12	55 ± 14	29 ± 9
BMI (kg/m ²) ^a	21 ± 3.4	27 ± 6.7	26 ± 5.0	20.8 ± 2.2
fasting glucose (mM) ^a	4.75 ± 0.43	4.90 ± 0.55	5.77 ± 0.99	5.32 ± 0.44
2-h glucose (mM) ^a	6.14 ± 0.93	9.17 ± 0.89	14.8 ± 4.24	7.27 ± 1.94

For the Ohashi dataset, data is given per glucose tolerance condition.

^aData presented as mean ± standard deviation.

Table 1. All subjects underwent a 75g-oral glucose tolerance test (OGTT) in the morning after an overnight fast. Fasting blood samples were drawn twice before oral ingestion of glucose. Blood samples were obtained at 10, 20, 30, 45, 60, 75, 90, 120, 150, 180, 210, 240 min after ingestion. Subjects remained at rest throughout the test. Blood samples were rapidly centrifuged⁴³.

Data preprocessing

Measurements of four subjects with missing values in the OGTT experiment were excluded from further analysis. The values reported in Table 1 are calculated on the data after exclusion. Unit conversions were performed to convert glucose from mgdL⁻¹ to mM and c-peptide from ngmL⁻¹ to nM. For the data from Fujita et al.⁴³, no measurements were dropped and the same unit conversions were applied, as with the data from Ohashi et al.

Differential equation model of c-peptide

The van Cauter model was used to describe the concentrations of c-peptide in the plasma and interstitial compartment²² (Fig. 2a). The original model, used to describe intravenously administered c-peptide, was extended to include endogenous production of c-peptide by the pancreas. The model equations for both compartments are given by:

$$\frac{dC^{\text{pl}}}{dt} = -(k_0 + k_2)C^{\text{pl}} + k_1 C^{\text{int}} + P(t) \quad (3)$$

$$\frac{dC^{\text{int}}}{dt} = k_2 C^{\text{pl}} - k_1 C^{\text{int}} \quad (4)$$

Where C^{pl} represents the concentration of c-peptide in the plasma compartment and C^{int} is the concentration of C-peptide in the interstitial compartment. Kinetic parameters k_0 – k_2 were calculated for each individual based on age, using equations provided by van Cauter et al.²², which are given as:

$$\begin{aligned} k_1 &= f \frac{\log(2)}{\tau_L} + (1-f) \frac{\log(2)}{\tau_S} \\ k_0 &= \frac{\log(2)}{\tau_S} \cdot \frac{\log(2)}{\tau_L \cdot k_1} \\ k_2 &= \frac{\log(2)}{\tau_S} + \frac{\log(2)}{\tau_L} - k_0 - k_1 \end{aligned}$$

For which the parameter values (f , τ_S , and τ_L) are given in Table 2 for the NGT, IGT, and T2DM groups.

Neural network component

The production of c-peptide $P(t)$ was modelled using a densely connected neural network with two inputs; the first was given by the difference in plasma glucose at time t compared to fasting values ($G_f(t) = G^{\text{pl}}(t) - G^{\text{pl}}(0)$) and the second was a learnable parameter β_i representing the inter-

Table 2 | Parameter values for computing the kinetic parameters for the van Cauter c-peptide model for the NGT, IGT and T2DM groups

Parameter (unit)	Value (NGT, IGT)	Value (T2DM)
f (-)	0.76	0.78
τ_S (min)	4.95	4.52
τ_L (min)	$29.2 + 0.14 \cdot \text{Age}$	

individual variability (fig. 2b). Plasma glucose values are obtained directly from the measured data using a forcing function. For timepoints in between measurements, glucose values are linearly interpolated.

The neural network contained two hidden layers each consisting of 4 nodes making use of tanh-activation functions, and an output layer of size 1, with a softplus activation function, resulting in 37 trainable weights.

The neural network architecture was selected through a grid search. Different architectures were obtained by varying the depth of the model between 1 and 2 layers, with layer sizes of 3–6 nodes, and 3 layers with layer sizes of 3 and 4 nodes. All models were trained on 70% of the train set and evaluated in the remaining 30%. The model that gave the lowest error for most individuals was selected. In case of a tie, the model with the lowest median error on all individuals was selected.

Initial conditions

For simulation, the whole system was assumed to be in steady state at $t = 0$, as subjects were fasting prior to the oral glucose tolerance test. The initial condition for plasma c-peptide (C^{pl}) was set to the measured fasting value at $t = 0$. For interstitial c-peptide, the initial condition was calculated using the steady-state assumption to be

$$C^{\text{int}}(t = 0) = \frac{k_2}{k_1} C^{\text{pl}}(t = 0) \quad (5)$$

Furthermore, to ensure the plasma c-peptide compartment was in steady-state at $t = 0$, the production term $P(t)$ including the neural network $N(G, \beta_i)$, describing c-peptide production was formulated as

$$P(t) = P_0 + N(G(t) - G(t = 0); \beta_i) - N(0; \beta_i) \quad (6)$$

Resulting in a production value at $t = 0$ of $P(t = 0) = P_0$, where P_0 was set as

$$k_0 C^{\text{pl}}(t = 0) \quad (7)$$

Parameter estimation

The neural network parameters were estimated on a randomly selected training subset containing 70% of the total samples, stratified according to the glucose tolerance condition. This training set was further divided into a true train set of 70% and a validation set of 30% of samples. This resulted in a true training set containing 49% of the entire dataset ($n = 57$), and a validation set containing 21% of the entire dataset ($n = 25$). Parameters were estimated on the true train set using the following loss function:

$$L_{\text{train}}(p_{\text{NN}}, \beta) = \sum_{i=1}^{N_{\text{train}}} \sum_{t=0}^T \left(C_{\text{model}}^{\text{pl}}(t|p_{\text{NN}}, \beta_i) - C_{\text{data},i}^{\text{pl}}(t) \right)^2 \quad (8)$$

Where p_{NN} are the parameters of the neural network, β represents the vector of all conditional parameters for each individual i out of the total of N_{train} individuals. Furthermore, $T = \{0, 30, 60, 90, 120\}$ represents the set of timepoints contained in the data. To prevent sign changes of β between individuals, $\log(\beta)$ was estimated, constraining β to the positive domain.

The parameter estimation for the universal differential equation models was then performed by sampling 25,000 initial candidate parameter sets and optimising the 25 candidate parameter sets that yielded the smallest initial objective function values. Subsequent optimization was performed using a two-stage optimizer, starting with Adam⁴⁴ for 1000 iterations with a learning rate of 10^{-2} . Starting from the endpoint of Adam, the LBFGS optimizer was used for a maximum of 1000 iterations or until convergence. Subsequently, for all trained 25 models, the neural network parameters were fixed and the conditional parameters were estimated on the validation set using the LBFGS optimizer, with the following loss function:

$$L_{\text{test}}(\beta_i) = \frac{|T|}{2} \ln(\sigma^2) + \frac{1}{2\sigma^2} \sum_{t=0}^T \left(C_{\text{model}}^{\text{pl}}(t|p_{\text{NN}}, \beta_i) - C_{\text{data},i}^{\text{pl}}(t) \right)^2 \quad (9)$$

The model that resulted in the lowest average loss function value in the individuals in the validation set was then selected as the best performing model.

After selection of the best performing model, the conditional parameters were reestimated on the full dataset, including the remaining 30% of the data that was not used until now ($n = 35$). Furthermore, for each individual, the variance in the residuals (σ_i^2) was estimated to enable the computation of confidence intervals on the conditional parameter (see “Identifiability analysis”). Estimation was performed using maximum likelihood estimation assuming zero-mean residuals, depicted in equation (10).

$$\text{NLL}(\beta_i, \sigma_i) = \frac{|T|}{2} \ln(\sigma_i^2) + \frac{1}{2\sigma_i^2} \sum_{t=0}^T \left(C_{\text{model}}^{\text{pl}}(t|p_{\text{NN}}, \beta_i) - C_{\text{data},i}^{\text{pl}}(t) \right)^2 \quad (10)$$

Identifiability analysis

To determine whether β_i was identifiable for each individual, we inspected the maximum likelihood function values (equation (10)) with the estimated σ_i^2 fixed, when varying β_i around its optimum. The 95% confidence interval of β_i was determined by the boundary values for the change in likelihood, defined in⁴⁵ to be $\Delta\text{NLL} \approx 7.16$. For an individual, β_i was defined as identifiable if these bounds were reached. If only one bound was reached, β_i was defined as practically unidentifiable. If neither bounds were reached, β_i was defined as unidentifiable⁴⁶.

Symbolic regression

For symbolic regression, initially 900 unique samples of the neural network output were created through combinations of 30 values for the conditional parameter β and incremental glucose values. Incremental glucose values were capped at zero, to reduce the complexity of the problem. Symbolic regression was then performed using the PySR package²⁵ using the settings listed in Table 3.

From the resulting equations, the top equation was selected using the ‘best’ option from the PySR package. This first selects all expressions with a loss smaller than at least 1.5 times the loss of the most accurate model. From these expressions, the model equation with the highest score is selected, defined as the negated derivative of the loss with respect to complexity²⁵.

The resulting equation was then simplified by amalgamating constants into a single learnable parameter. As the incremental glucose values used to train the symbolic equation were capped at zero, production was set at a value of zero when $G^{\text{pl}}(t) < G^{\text{pl}}(0)$.

Programming

Both the ordinary differential equation models, as well as the universal differential equation models used in this research were implemented in the Julia programming language, using the ‘OrdinaryDiffEq.jl’ package⁴⁷.

Table 3 | Settings for the symbolic regression algorithm from PySR

Parameter	Value
Iterations	25,000
Binary operators	[+, *]
Unary operators	inv(x) = 1/x
Processes	6
Maximum size	18
Populations	24

Data availability

All code and data used to produce the results and analyses can be found in the GitHub repository linked to this publication: <https://github.com/Computational-Biology-TUe/conditional-ude>.

Received: 5 March 2025; Accepted: 19 July 2025;
Published online: 31 July 2025

References

1. Nielsen, J. Systems biology of metabolism: a driver for developing personalized and precision medicine. *Cell Metab.* **25**, 572–579 (2017).

2. Wysham, C. & Shubbrook, J. Beta-cell failure in type 2 diabetes: mechanisms, markers, and clinical implications. *Postgrad. Med.* **132**, 676–686 (2020).

3. Jones, A. G. et al. Markers of β -cell failure predict poor glycemic response to GLP-1 receptor agonist therapy in type 2 diabetes. *Diab. Care* **39**, 250–257 (2016).

4. Thong, K. Y. et al. The association between postprandial urinary c-peptide creatinine ratio and the treatment response to liraglutide: a multi-centre observational study. *Diabet. Med.* **31**, 403–411 (2014).

5. Ashley, E. A. Towards precision medicine. *Nat. Rev. Genet.* **17**, 507–522 (2016).

6. Coral, D. E. et al. Subclassification of obesity for precision prediction of cardiometabolic diseases. *Nat. Med.* <https://doi.org/10.1038/s41591-024-03299-7> (2024).

7. Berger, M. F. & Mardis, E. R. The emerging clinical relevance of genomics in cancer medicine. *Nat. Rev. Clin. Oncol.* **15**, 353–365 (2018).

8. Swanson, K., Wu, E., Zhang, A., Alizadeh, A. A. & Zou, J. From patterns to patients: advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. *Cell* **186**, 1772–1791 (2023).

9. Finotello, F. & Eduati, F. Multi-omics profiling of the tumor microenvironment: paving the way to precision immuno-oncology. *Front. Oncol.* **8**, 430 (2018).

10. Wang, H. et al. Deep learning in systems medicine. *Brief. Bioinform.* **22**, 1543–1559 (2021).

11. Ching, T. et al. Opportunities and obstacles for deep learning in biology and medicine. *J. R. Soc. Interface* **15**, 20170387 (2018).

12. Sapoval, N. et al. Current progress and open challenges for applying deep learning across the biosciences. *Nat. Commun.* **2022** 13:1 **13**, 1–12 (2022).

13. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **1**, 206–215 (2019).

14. Lauzeral, N. et al. A model order reduction approach to create patient-specific mechanical models of human liver in computational medicine applications. *Comput. Methods Prog. Biomed.* **170**, 95–106 (2019).

15. Kuepfer, L. & Schuppert, A. Systems medicine in pharmaceutical research and development. *Methods Mol. Biol.* **1386**, 87–104 (2016).

16. O'Donovan, S. D. et al. Quantifying the effect of nutritional interventions on metabolic resilience using personalized computational models. *iScience* **27**, 109362 (2024).

17. Erdős, B. et al. Quantifying postprandial glucose responses using a hybrid modeling approach: combining mechanistic and data-driven models in the Maastricht study. *PLoS ONE* **18**, e0285820 (2023).

18. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).

19. Rackauckas, C. et al. Universal differential equations for scientific machine learning. *ArXiv* (2021).

20. de Rooij, M., Erdős, B., van Riel, N. A. W. & O'Donovan, S. D. Physiology-informed regularisation enables training of universal differential equation systems for biological applications. *PLOS Comput. Biol.* **21**, e1012198 (2025).

21. Philipps, M., Körner, A., Vanhoefer, J., Pathirana, D. & Hasenauer, J. Non-Negative Universal Differential Equations With Applications in Systems Biology. *IFAC-PapersOnLine* **58**, 25–30 (2024).

22. van Cauter, E., Mestrez, F., Sturis, J. & Polonsky, K. S. Estimation of insulin secretion rates from C-peptide levels. Comparison of individual and standard kinetic parameters for C-peptide clearance. *Diabetes* **41**, 368–377 (1992).

23. Okuno, Y. et al. Postprandial serum c-peptide to plasma glucose concentration ratio correlates with oral glucose tolerance test- and glucose clamp-based disposition indexes. *Metab. Clin. Exp.* **62**, 1470–1476 (2013).

24. Ohashi, K. et al. Glucose homeostatic law: Insulin clearance predicts the progression of glucose intolerance in humans. *PLoS ONE* **10**, e0143880 (2015).

25. Cranmer, M. Interpretable machine learning for science with PySR and symbolicregression.jl. *ArXiv*. <https://arxiv.org/abs/2305.01582v3> (2023).

26. Holst, J. J., Gasbjerg, L. S. & Rosenkilde, M. M. The role of incretins on insulin function and glucose homeostasis. *Endocrinology* **162**, bqab065 (2021).

27. Pasveer, Y. M. et al. Does GLP-1 cause post-bariatric hypoglycemia: 'computer says no'. *Comput. Methods Prog. Biomed.* **257**, 108424 (2024).

28. Henquin, J. C. Triggering and amplifying pathways of regulation of insulin secretion by glucose. *Diabetes* **49**, 1751–1760 (2000).

29. Aguayo-Mazzucato, C. Functional changes in beta cells during ageing and senescence. *Diabetologia* **63**, 2022–2029 (2020).

30. de Lepper, A. G. W. et al. From evidence-based medicine to digital twin technology for predicting ventricular tachycardia in ischaemic cardiomyopathy. *J. R. Soc. Interface* **19**, 20220317 (2022).

31. Zeevi, D. et al. Personalized nutrition by prediction of glycemic responses. *Cell* **163**, 1079–1094 (2015).

32. Lutsker, G. et al. From glucose patterns to health outcomes: a generalizable foundation model for continuous glucose monitor data analysis. *ArXiv* (2024).

33. Maas, A. H. et al. A physiology-based model describing heterogeneity in glucose metabolism: The core of the eindhoven diabetes education simulator (e-des). *J. Diab. Sci. Technol.* **9**, 282–292 (2015).

34. Ha, J. et al. Estimating insulin sensitivity and β -cell function from the oral glucose tolerance test: validation of a new insulin sensitivity and secretion (ISS) model. *Am. J. Physiol. Endocrinol. Metab.* **326**, E454–E471 (2024).

35. Hovorka, R., Chassin, L., Luzio, S. D., Playle, R. & Owens, D. R. Pancreatic beta-cell responsiveness during meal tolerance test: model assessment in normal subjects and subjects with newly diagnosed noninsulin-dependent diabetes mellitus. *J. Clin. Endocrinol. Metab.* **83**, 744–750 (1998).

36. Topp, B., Promislow, K., Devries, G., Miura, R. M. & Finegood, D. T. A model of β -cell mass, insulin, and glucose kinetics: pathways to diabetes. *J. Theor. Biol.* **206**, 605–619 (2000).

37. Møller, J. B. et al. Ethnic differences in insulin sensitivity, β -cell function, and hepatic extraction between Japanese and caucasians: a

npj Systems Biology and Applications | (2025)11:84

9

- minimal model analysis. *J. Clin. Endocrinol. Metab.* **99**, 4273–4280 (2014).
38. van Sloun, B. et al. The impact of amino acids on postprandial glucose and insulin kinetics in humans: a quantitative overview. *Nutrients* **12**, 3211 (2020).
 39. Schmid, N., Fernandes del Pozo, D., Waegeman, W. & Hasenauer, J. Assessment of uncertainty quantification in universal differential equations. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **383**, 20240444 (2025). Publisher: Royal Society.
 40. Tornøe, C. W., Agersø, H., Jonsson, E. N., Madsen, H. & Nielsen, H. A. Non-linear mixed-effects pharmacokinetic/pharmacodynamic modelling in NLME using differential equations. *Comput. Methods Prog. Biomed.* **76**, 31–40 (2004).
 41. Martensen, C. J., Korsbo, N., Ivaturi, V. & Sager, S. Data-driven discovery of feedback mechanisms in acute myeloid leukaemia: alternatives to classical models using deep nonlinear mixed effect modeling and symbolic regression. *bioRxiv*. <https://www.biorxiv.org/content/10.1101/2024.06.17.599366v1.abstract> (2024).
 42. Ohashi, K. et al. Increase in hepatic and decrease in peripheral insulin clearance characterize abnormal temporal patterns of serum insulin in diabetic subjects. *npj Syst. Biol. Appl.* **4**, 1–12 (2018).
 43. Fujita, S. et al. Four features of temporal patterns characterize similarity among individuals and molecules by glucose ingestion in humans. *npj Syst. Biol. Appl.* **2022** 8:1 **8**, 1–16 (2022).
 44. Kingma, D. P. & Ba, J. L. Adam: A method for stochastic optimization. in *Proceeding 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* (2014).
 45. Tönsing, C., Steiert, B., Timmer, J. & Kreutz, C. Likelihood-ratio test statistic for the finite-sample case in nonlinear ordinary differential equation models. *PLoS Comput. Biol.* **19**, e1011417 (2023).
 46. Raue, A. et al. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* **25**, 1923–1929 (2009).
 47. Rackauckas, C. & Nie, Q. Differentialequations.jl—a performant and feature-rich ecosystem for solving differential equations in Julia. *J. Open Res. Softw.* **5**, 15 (2017).

Acknowledgements

The research presented in this manuscript was supported by a Starting Package from the Eindhoven AI Systems Institute (EASI) awarded to S.O'D. N.A.W.v.R is supported by a grant from the Dutch Research Council (NWO) [<https://www.nwo.nl/>] as part of the Diagame project (project number 645.001.003). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. All other authors did not receive specific funding.

Author contributions

M.d.R.: Conceptualisation, Methodology, Software, Investigation, Formal analysis, Writing - original draft, Visualisation. N.A.W.v.R.:

Conceptualisation, Resources, Writing - Review & Editing, Supervision. S.D.O'D.: Conceptualisation, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Competing interests

The authors declare no competing interests.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used 'Writefull AI' in order to suggest language improvements to the initial draft. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41540-025-00570-6>.

Correspondence and requests for materials should be addressed to Max de Rooij.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025