

Cell Reports Methods, Volume 3

Supplemental information

**Uncovering personalized glucose responses
and circadian rhythms from multiple wearable
biosensors with Bayesian dynamical modeling**

Nicholas E. Phillips, Tinh-Hai Collet, and Felix Naef

1 Supplementary Tables

Study ID	MSF _{SC} (hrs)	IPAQ (METs-hrs/wk)	PSQI	App adherence (%)	Valid Actiheart (%)
01	4.8	13.65	7	100.0	72.5
02	5.5	9.65	2	100.0	58.6
03	3.3	5.30	1	100.0	59.7
04	4.9	15.30	7	100.0	80.5
05	5.0	37.65	3	100.0	85.5
06	5.3	9.49	4	100.0	73.3
07	3.3	35.30	4	100.0	Not used
08	4.2	24.95	5	100.0	70.3
09	3.7	2.47	7	100.0	93.7
10	4.6	53.10	7	100.0	90.8
11	3.6	3.10	3	100.0	72.4
12	5.6	28.95	2	100.0	94.4
13	4.5	14.47	9	100.0	83.4
14	2.3	10.95	2	100.0	74.3
15	4.7	23.30	5	100.0	41.4
16	3.6	39.30	5	100.0	89.6
17	4.5	5.00	5	83.3	82.6
18	5.0	25.10	4	100.0	85.6
19	3.8	18.60	4	100.0	91.3
20	4.2	9.60	8	100.0	89.2
21	4.5	14.16	5	100.0	60.9
22	5.8	3.22	6	Not used	81.3
23	4.3	24.95	7	100.0	86.3
24	2.5	28.82	5	100.0	74.6
25	4.2	9.65	5	100.0	91.8

Table S1: Chronotype and data quality metrics for study participants. Related to STAR Methods. MSF_{SC} (hours): chronotype variable calculated from the Munich Chronotype Questionnaire (MCTQ); IPAQ: physical activity as measured with the International Physical Activity Questionnaire (IPAQ); PSQI: sleep quality as assessed with the Pittsburgh Sleep Quality Index. Range 0-27, where higher scores reflect lower quality sleep.; App adherence (%): the percentage of days with two recorded meals separated by at least 5 hours, calculated using all valid study days and excluding the first and last study days; Valid Actiheart (%): the percentage of Actiheart data that passes quality filtering for each participant.

2 Supplementary Figures

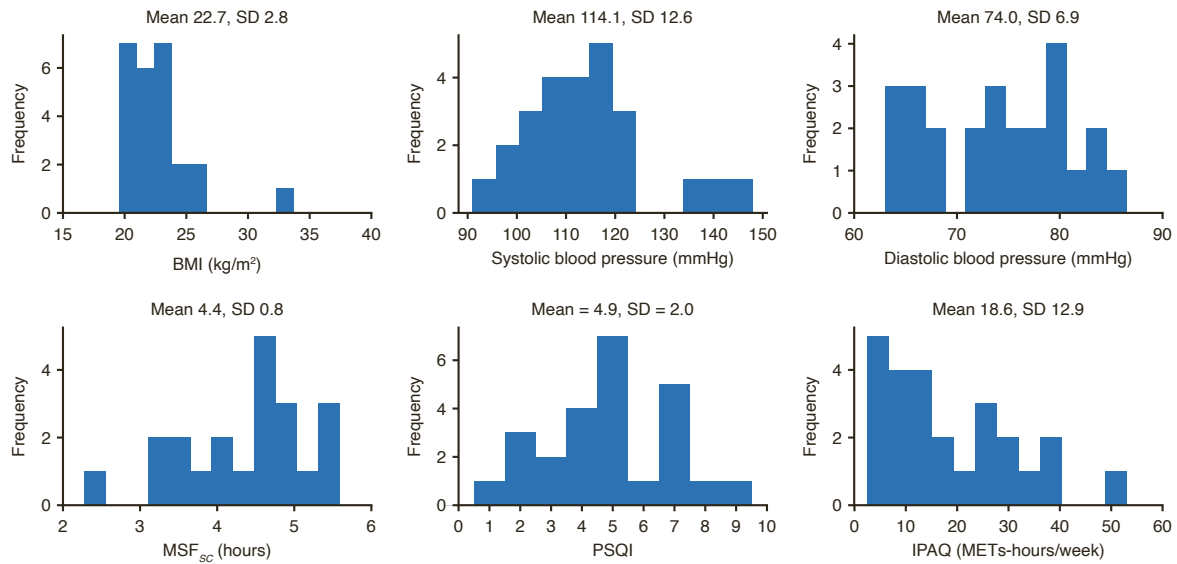


Figure S1: The distribution of participant characteristics in the study population, displayed as histograms. Related to STAR Methods. BMI: body mass index. MSF_{sc} (hours): chronotype variable calculated from the Munich Chronotype Questionnaire (MCTQ); PSQI: sleep quality as assessed with the Pittsburgh Sleep Quality Index. Range 0-27, where higher scores reflect lower quality sleep; IPAQ: physical activity as measured with the International Physical Activity Questionnaire (IPAQ).

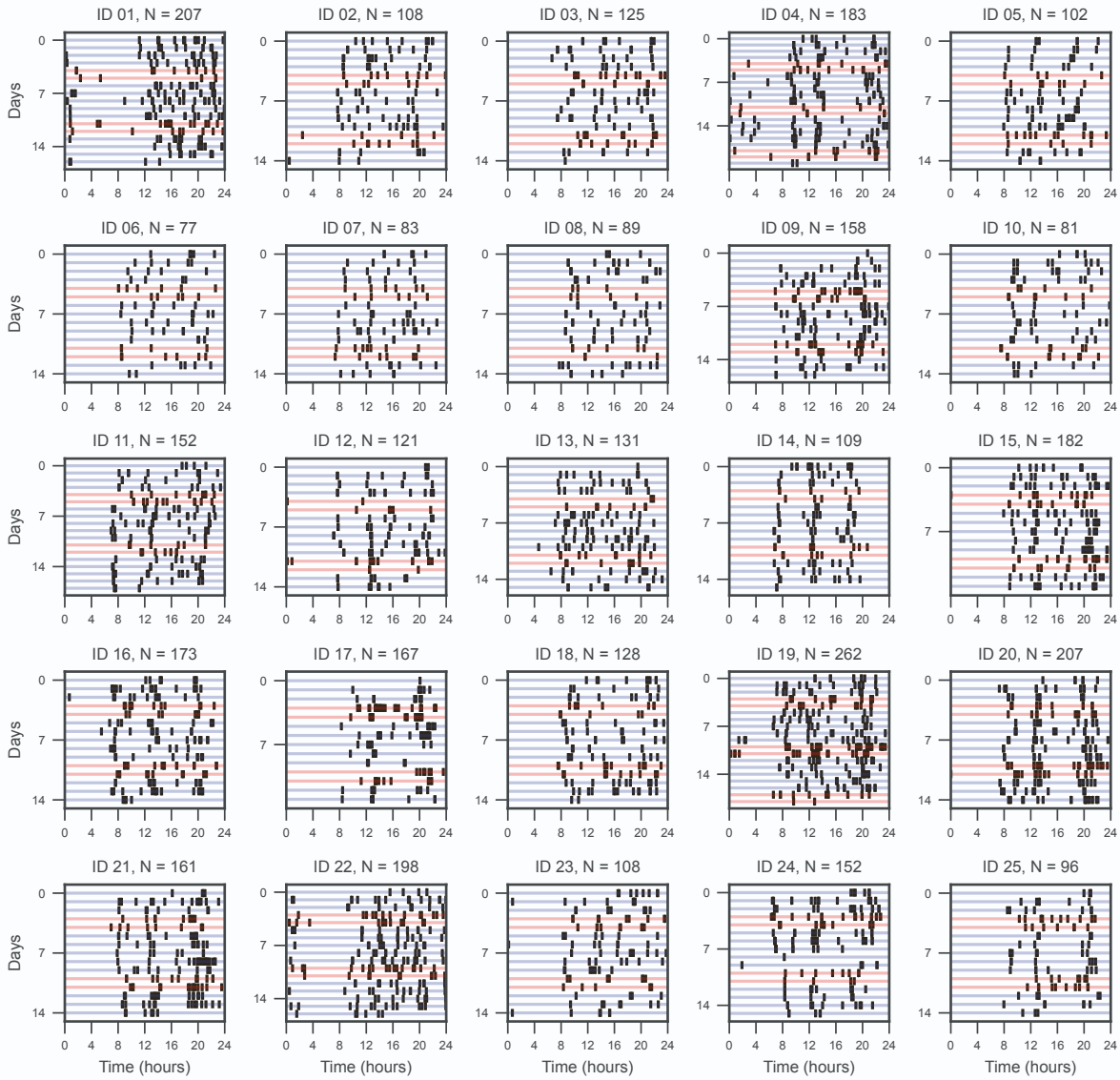


Figure S2: All recorded food and drink events for all study participants. Related to Figure 1. Each horizontal line represents a different study day, and food and drink events are marked as vertical lines. Day 0 represents the first study day. Blue horizontal lines denote weekdays while red horizontal lines represent the weekend. The total number of recorded ingestion events is indicated with N.

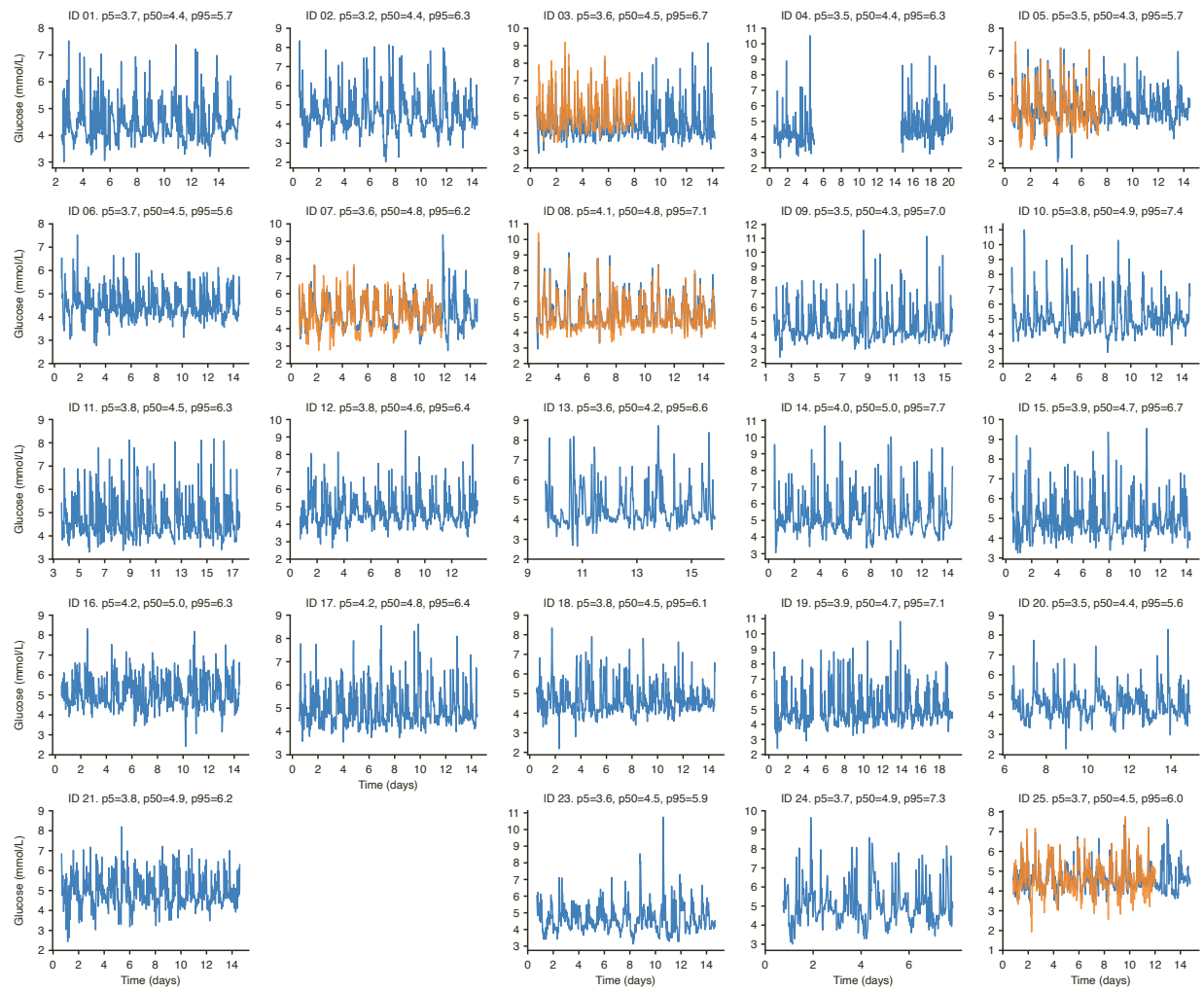


Figure S3: CGM data for all study participants. Related to Figure 1. For each study participant, the 5th, 50th and 95th percentiles are indicated along with the study ID. For five participants there is a second sensor (orange line). Device failure for ID 22 meant that the data is missing for this participant, and an empty space is left.

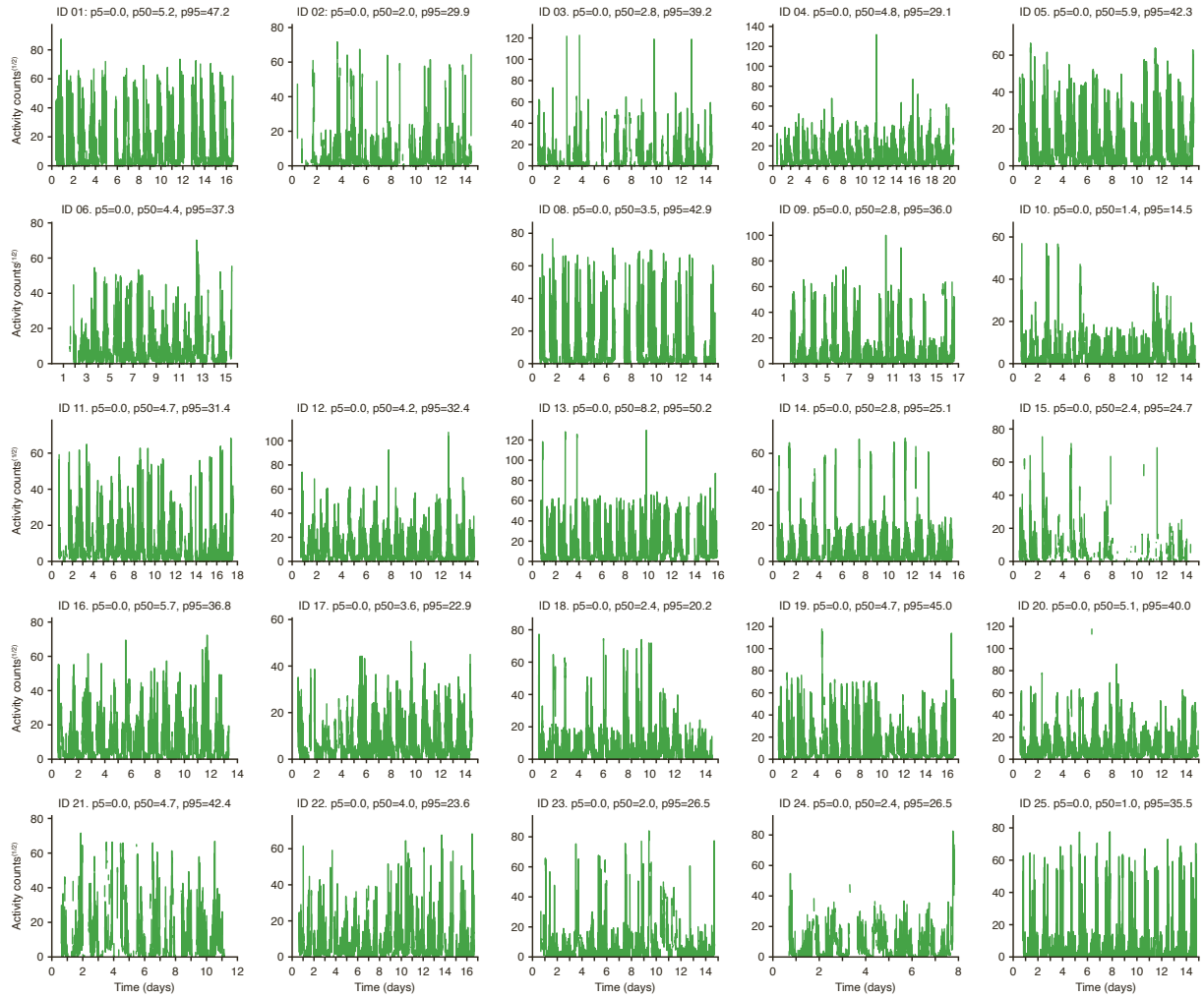


Figure S4: Activity data for all study participants, as measured with the Actiheart device. Related to Figure 1. For each study participant, the 5th, 50th and 95th percentiles are indicated along with the study ID. Device failure for ID 07 meant that the data is missing for this participant, and an empty space is left.

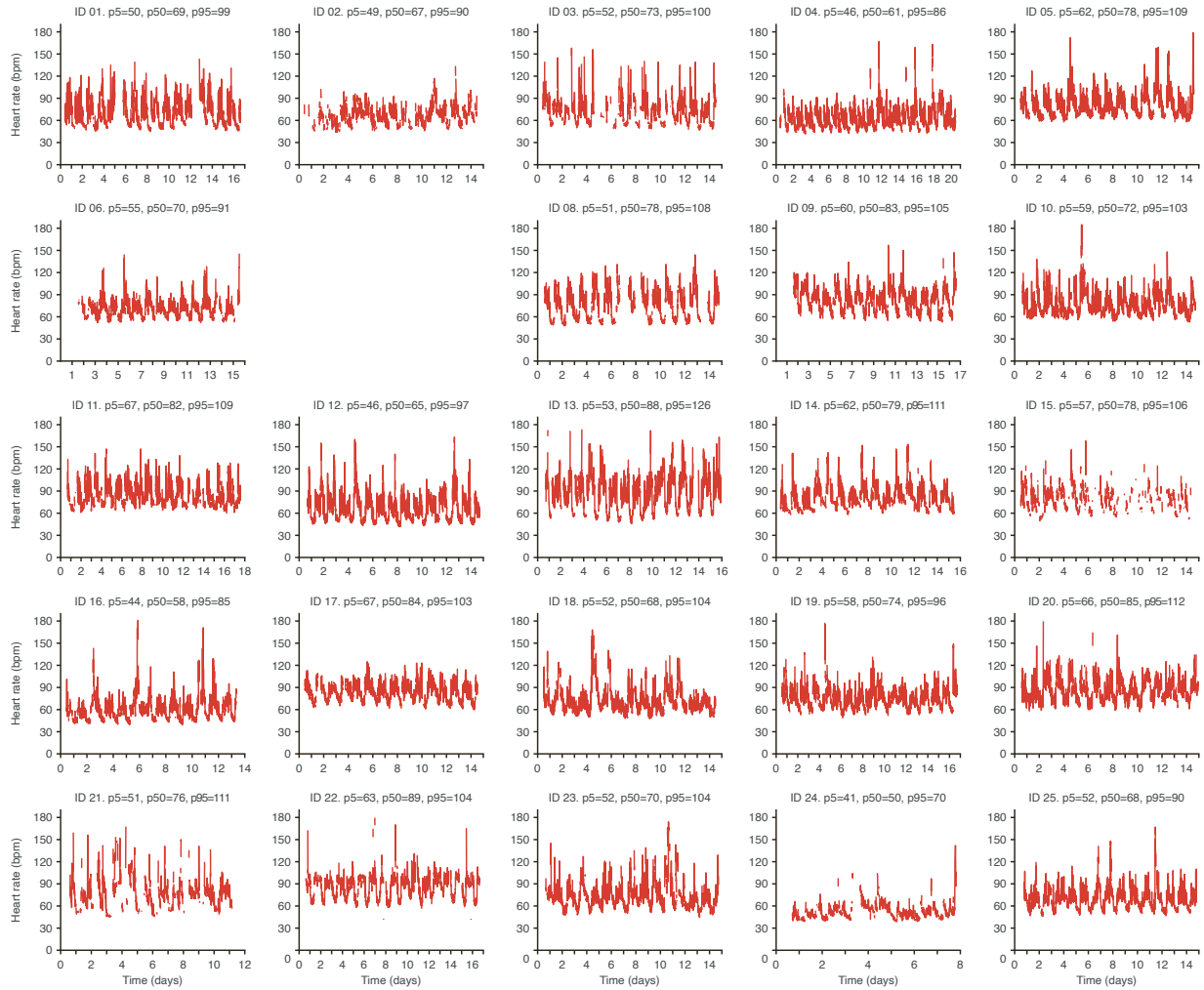


Figure S5: Heart rate data for all study participants (bpm: beats per minute), as measured with the Actiheart device. Related to Figure 1. For each study participant, the 5th, 50th and 95th percentiles are indicated along with the study ID. Device failure for ID 07 meant that the data is missing for this participant, and an empty space is left.



Figure S6: Heart rate variability data for all study participants (RMSSD: root mean square of successive differences between heartbeats), as measured with the Actiheart device. Related to Figure 1. For each study participant, the 5, 50 and 95% percentiles are indicated along with the study ID. Device failure for ID 07 meant that the data is missing for this participant, and an empty space is left.

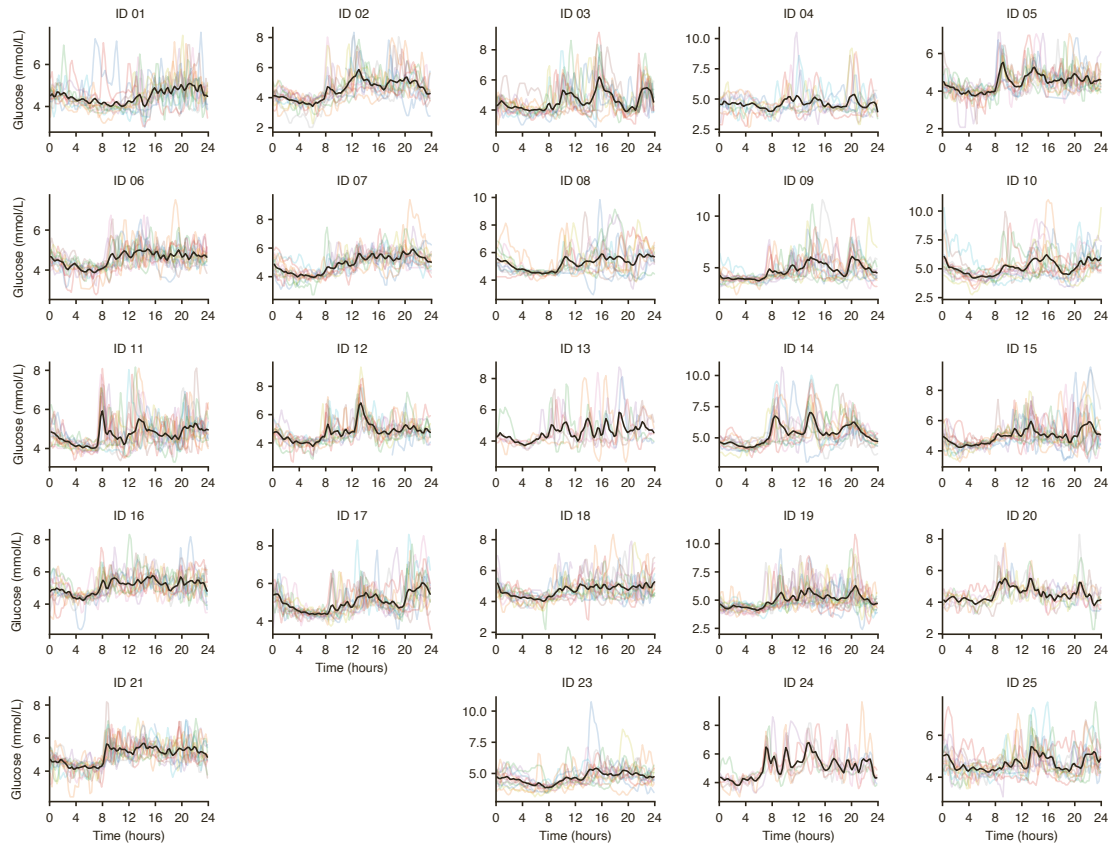


Figure S7: Glucose time-series data over a two-week period for all participants. Related to Figure 1. Individual days are represented by different coloured lines and the mean glucose level over all measured days in black.

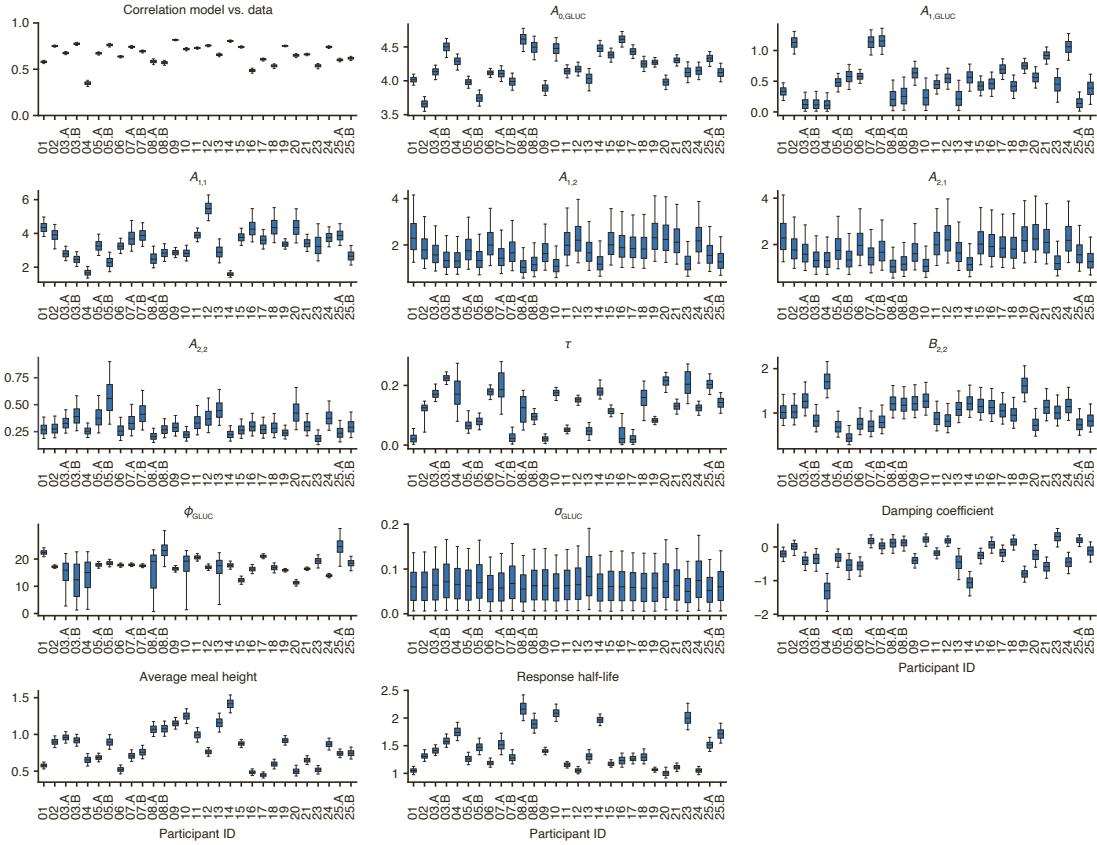


Figure S8: Posterior distributions (shown as boxplots) for all glucose model parameters (Model 1) across all participants. Related to Figure 3. The first panel shows the correlation between the inferred model (including meals and circadian time) and the glucose CGM data for each participant. The boxplots represent the 25th, median (50th) and 75th percentiles of the posterior distribution and the whiskers represent the 5th and 95th percentiles.

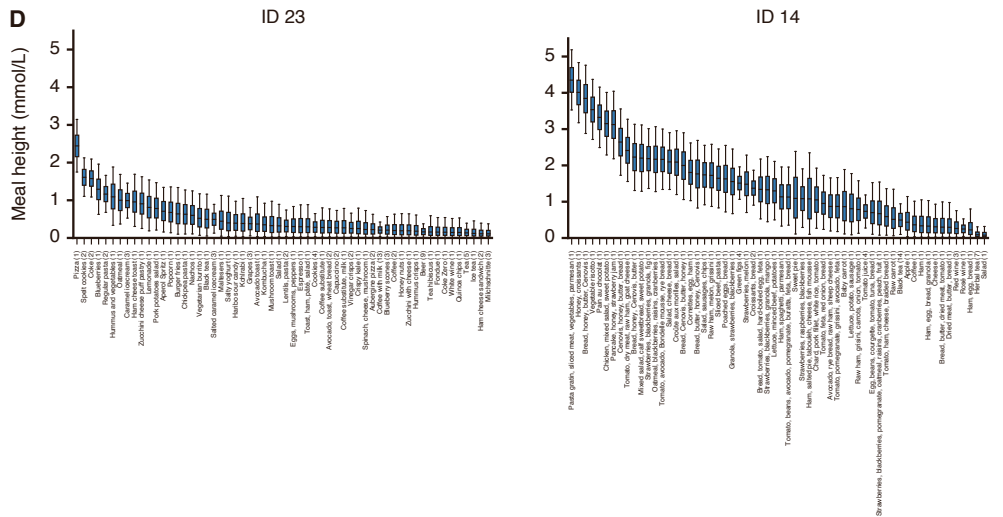
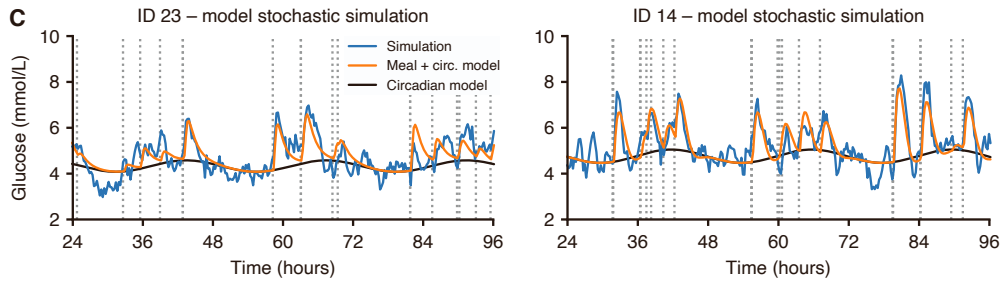
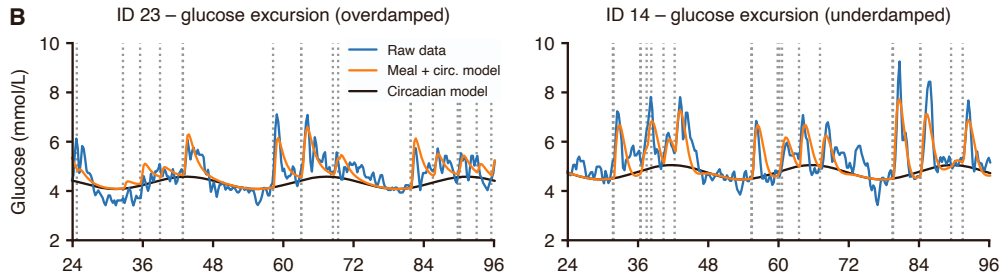
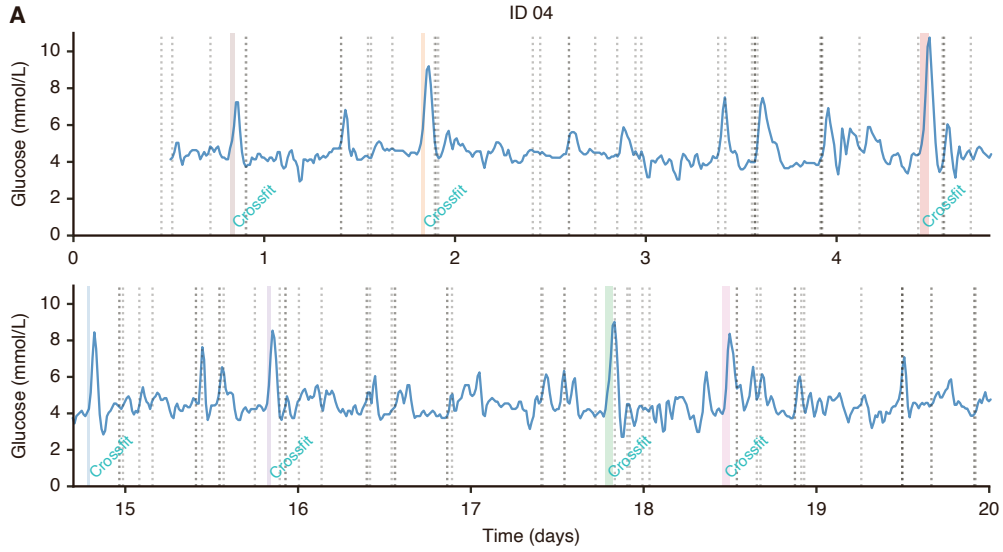


Figure S9: Accessory figures related to Figure 3. (A) The relationship between glucose levels (blue), meals (grey dotted lines) and sport events (shaded areas with annotations) for participant ID 04. Due to a sensor failure, the glucose data for ID 04 is in two blocks. (B-C) Simulating glucose traces using the full model with random fluctuations creates realistic glucose traces. (B) Examples comparing the CGM data (blue) with the model prediction incorporating circadian dynamics (black) plus meal consumption (orange) for two participants with overdamped and underdamped dynamics, respectively. Same panels as Figure 3E-F, repeated here for easy comparison with panel C. (C) Simulations from the full glucose model that also includes random fluctuations (blue). The timestamps of meals are shown as black dashed lines. (D) Posterior distribution over inferred meal heights for ID 14 and 23, where meal heights are defined as glucose increases relative the baseline glucose levels. The number indicates the number of times the meal was consumed. The boxplots represent the 25th, median (50th) and 75th percentiles of the posterior distribution and the whiskers represent the 5th and 95th percentiles.

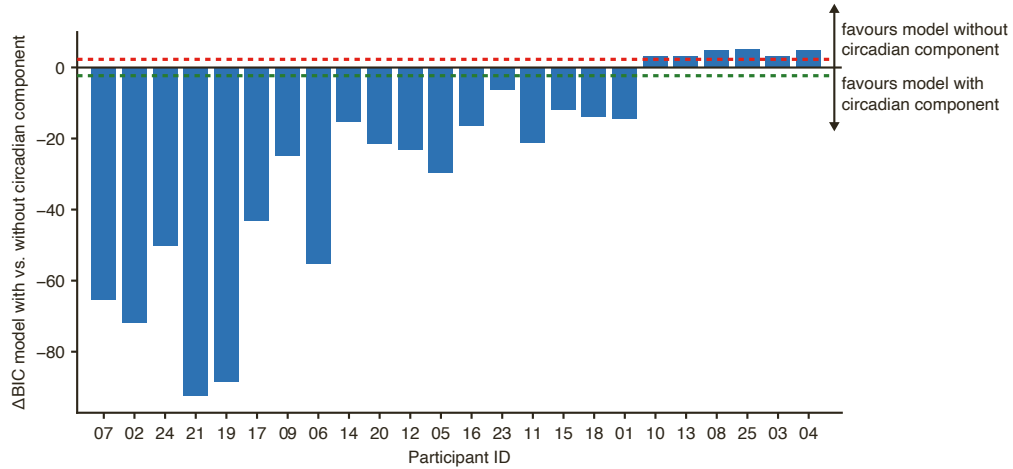


Figure S10: The difference in BIC score between Model 1 with a circadian component and Model 1 without a circadian component. Related to Figure 4. The two models were first optimised to their maximum a posteriori probability (MAP) estimates, then the BIC score was calculated. A negative difference in BIC score favours Model 1 with a circadian component, while a positive difference in BIC score favours the Model 1 without a circadian component. The dotted red line marks the threshold at which Model 1 without the circadian component is preferred, while the dotted green line marks the threshold at which Model 1 with the circadian component is preferred.

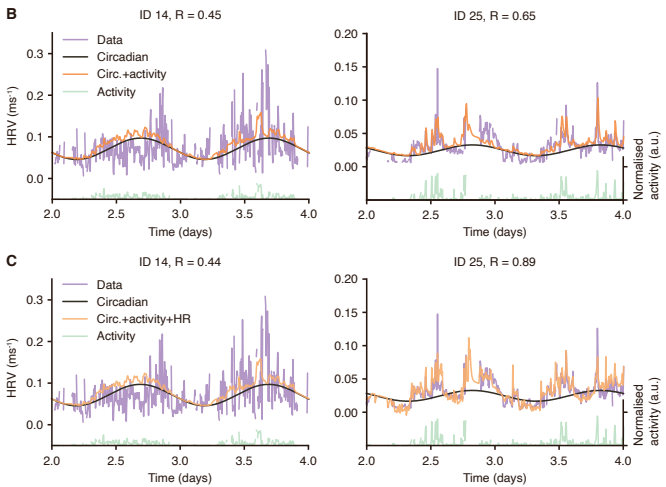
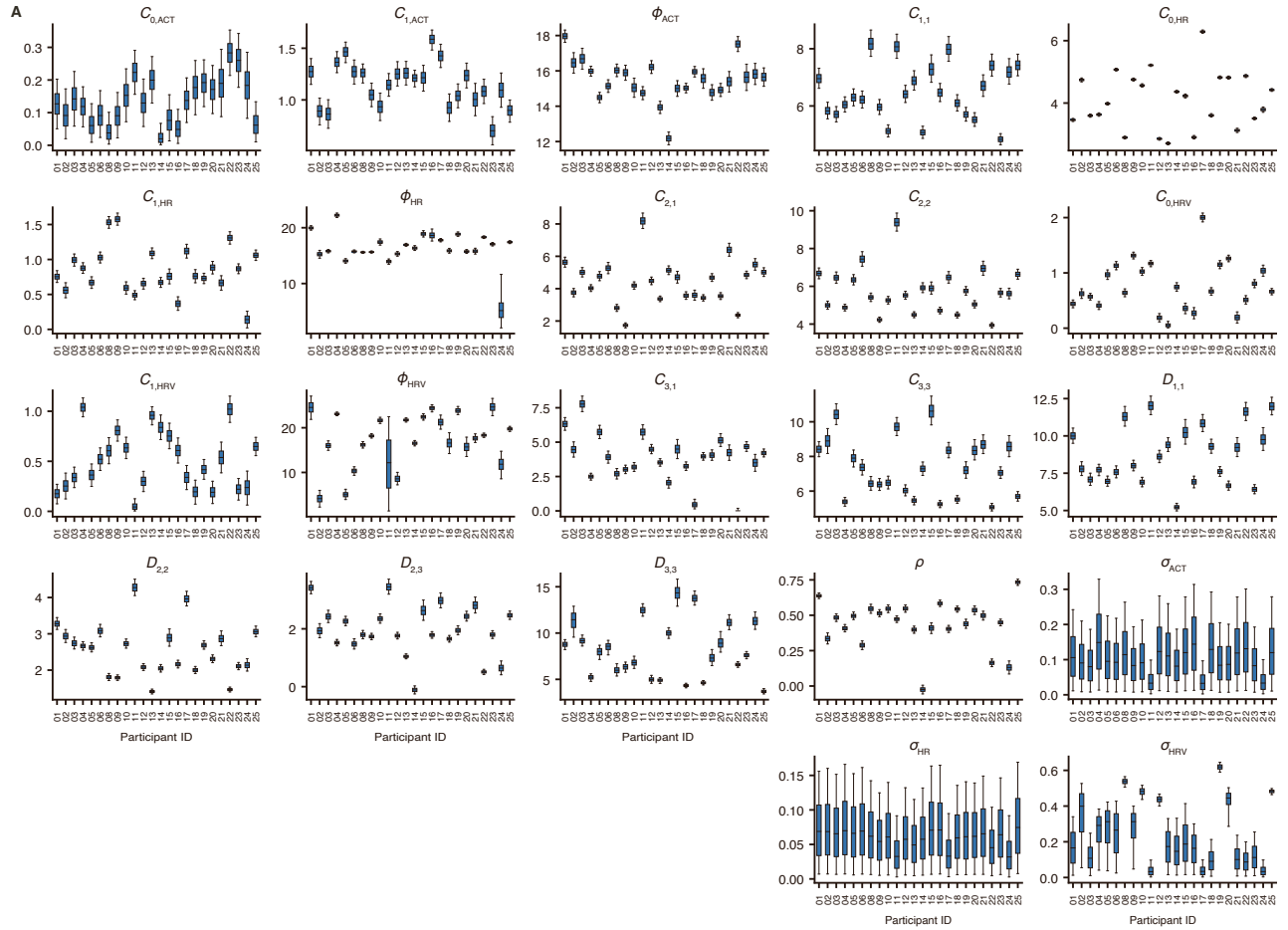


Figure S11: Related to Model 2 and Figure 5. (A) Posterior distributions (shown as boxplots) for all physical and heart activity model parameters (Model 2) across all participants. The boxplots represent the 25th, median (50th) and 75th percentiles of the posterior distribution and the whiskers represent the 5th and 95th percentiles. (B-C) Examples demonstrating the improvement in prediction quality when accounting for HR in predictions of HRV. (B) The prediction of HRV using only circadian time and physical activity. Blue: HRV data; black: baseline circadian rhythm; green: physical activity (shown on normalised scale where 1 represents the maximum value); orange: model prediction using circadian rhythms and physical activity. (C) The prediction of HRV using circadian time, physical activity and HR. Blue: HRV data; black: baseline circadian rhythm; green: physical activity (shown on normalised scale where 1 represents the maximum value); orange: model prediction using circadian rhythms, physical activity and HR.

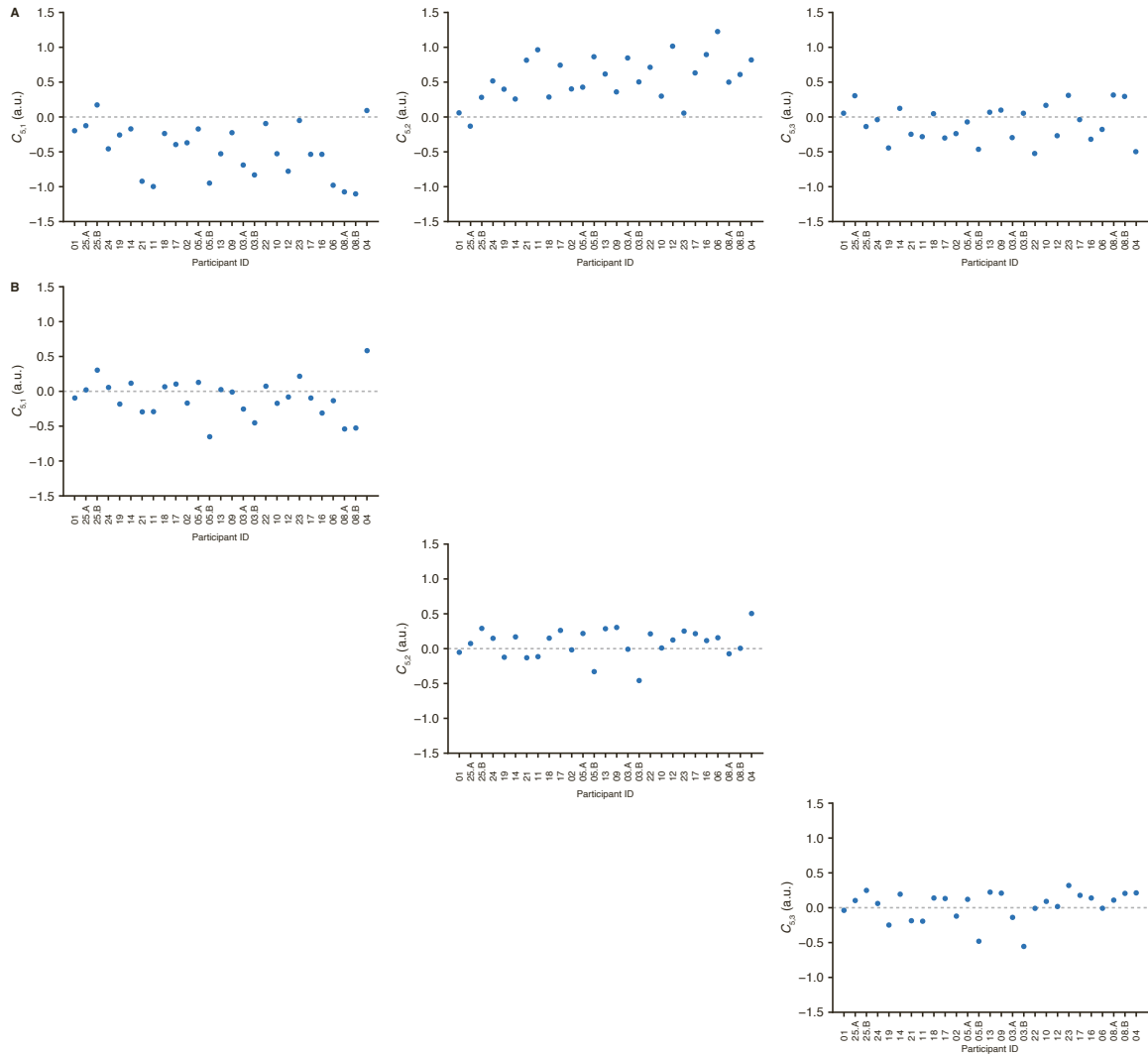


Figure S12: Model parameters when only one input from the physical and heart activity model is allowed. Related to Figure 5. (A) Maximum a posteriori probability (MAP) estimate of coefficients C_{51} , C_{51} and C_{51} with all physical and heart activity inputs. (B) MAP estimate of the coefficients C_{51} when physical activity (via parameter C_{51}) is the only physical and heart activity input into the glucose model. (C) MAP estimate of the coefficients C_{52} when heart rate (via parameter C_{52}) is the only physical and heart activity input into the glucose model. (D) MAP estimate of the coefficients C_{53} when heart rate variability (via parameter C_{53}) is the only physical and heart activity input into the glucose model.

3 Methods S1: details of the Kalman filter model, related to STAR Methods

3.1 Missing data and data quality

For each participant, we collected data for two weeks using the following devices: 1) timestamps of food/drinks and text annotations collected with the smartphone application (app) MyCircadianClock; 2) continuous glucose monitoring (CGM) using the FreeStyle Libre Pro device (Abbott); 3) physical activity, heart rate (HR) and heart rate variability (HRV using RMSSD^{-1}) using the Actiheart device (CamNTEch). FreeStyle CGMs were replaced when devices fell off participants within the first week, which occurred 6 times, and data was not used when there was less than 48 hours of valid measurements. For one participant (ID 22), the FreeStyle CGM device failed without falling off, but due to device blinding this was not detected until the end of the experiment, and hence there is no glucose data for this participant. While the Actiheart device is waterproof, participants were permitted to briefly remove the device during showers and baths. At each time point the Actiheart produces an estimated quality of the signal (range 0-1), and we filtered the data based on a threshold of 0.8. After filtering the signal based on the quality, an average of 78.5% of the data was included, although this was specific to each participant (Table S1). There was an Actiheart device failure for one participant (ID 07). The Actiheart data was exported with a time resolution of 5 minutes, while the FreeStyle CGM device records glucose every 15 minutes. To align the Actiheart and FreeStyle CGM time series, we chose the Actiheart measurement that was closest in time to the corresponding CGM measurement. This induces a maximum misalignment of 2.5 minutes, which we consider acceptable given that glucose typically has a response time >1 hour (Figure 3A).

3.2 Modelling wearable time series data

We use a linear Gaussian state space model (otherwise known as a Kalman filter, see Särkkä and Solin [S1] for details) to analyse the time series generated by the wearable devices, and here we will first outline the general parameter inference framework before giving details for the specific models used. To account for both dynamics in underlying physiological variables changes as well as measurement error of the devices, this approach considers underlying physiological variables (referred to as latent state variables) $\mathbf{x}(t_k)$ (measured at times t_k) and observed variables \mathbf{y}_k . The linear Gaussian state space model of the wearable data has the following three components:

3.2.1 Dynamic model

The dynamic model $\mathbf{x}(t_k) \sim p(\mathbf{x}(t_k) | \mathbf{x}(t_{k-1}))$ describes the transition probabilities of the underlying physiological variables. Our starting point is a continuous-time system of stochastic differential equations (SDEs)

$$d\mathbf{x}(t) = W\mathbf{x}(t)dt + d\boldsymbol{\beta}, \quad (1)$$

where W is a matrix describing the interactions between the variables and $\boldsymbol{\beta}$ is a brownian noise term with covariance matrix Q . The specific forms of W and Q are unique for each model and will be described in the following sections. We then convert this system of continuous-time SDEs into a model where time is discrete.

$$\mathbf{x}(t_k) = F_{k-1}\mathbf{x}(t_{k-1}) + N(0, \Sigma_{k-1}), \quad (2)$$

where F_k is the state-transition model and Σ_k is the covariance of the process noise. To perform this time discretisation we use the following two formulas, where details on their precise calculation for each model will also be specified below

$$\begin{aligned} F_k &= \exp(W\Delta t_k), \\ \Sigma_k &= \int_0^{\Delta t_k} \exp(W(\Delta t_k - \tau))Q \exp(W(\Delta t_k - \tau))^\top d\tau. \end{aligned} \quad (3)$$

3.2.2 Initial distribution

The initial distribution $\mathbf{x}(t_0) \sim p(\mathbf{x}(t_0))$ specifies the distribution of the physiological variables at the initial time $k = 0$. This is specified with $p(\mathbf{x}(t_0)) = N(\mathbf{x}(t_0) | \mathbf{m}_0, P_0)$, and where we use the stationary distribution of the SDE to define the mean and covariance \mathbf{m}_0 and P_0 .

3.2.3 Measurement model

The measurement model $\mathbf{y}_k \sim p(\mathbf{y}_k | \mathbf{x}(t_k))$ describes the observation process, which again assumes a Gaussian distribution.

$$\mathbf{y}(t_k) = H_k\mathbf{x}(t_k) + N(\mathbf{m}_k, R_k), \quad (4)$$

where H_k is the observation matrix and \mathbf{m}_k and R_k represents the mean and covariance of the observation noise, respectively.

3.3 The marginal posterior distribution of parameters

The parameters of the model can be represented with the parameter vector $\boldsymbol{\theta}$, and the goal is to infer the parameters for each participant. Within a Bayesian framework, the parameters of the model can be estimated from the data as follows

$$p(\boldsymbol{\theta} | \mathbf{y}_{1:T}) \propto p(\boldsymbol{\theta})p(\mathbf{y}_{1:T} | \boldsymbol{\theta}), \quad (5)$$

where $p(\boldsymbol{\theta})$ is the prior distribution of parameters and $p(\mathbf{y}_{1:T} | \boldsymbol{\theta})$ is the likelihood of observing the temporal data $\mathbf{y}_{1:T}$ given the set of parameters $\boldsymbol{\theta}$. Considering the time series sequence of data, the likelihood term for a given set of parameters $\boldsymbol{\theta}$ can be re-expressed as

$$p(\mathbf{y}_{1:T} | \boldsymbol{\theta}) = p(\mathbf{y}_1 | \boldsymbol{\theta}) \prod_{k=2}^T p(\mathbf{y}_k | \mathbf{y}_{1:k-1}, \boldsymbol{\theta}), \quad (6)$$

and the sequence of distributions $p(\mathbf{y}_k | \mathbf{y}_{1:k-1}, \boldsymbol{\theta})$ are calculated within a Kalman filtering framework that we will now describe.

3.4 Kalman Filter

We will now describe the Bayesian filtering approach used to estimate the likelihood term $p(\mathbf{y}_{1:T} | \boldsymbol{\theta})$. Starting with the mean \mathbf{m}_0 and covariance P_0 of the initial distribution $p(\mathbf{x}(t_0)) = \text{N}(\mathbf{x}(t_0) | \mathbf{m}_0, P_0)$, the Kalman filter recursion operates from the first to last data point $k = 1, 2, \dots, T$. At each time point, there is first a prediction step that predicts the current state of the latent variables $\mathbf{x}(t_k)$ based on all previous measurements up to \mathbf{y}_{k-1}

$$p(\mathbf{x}(t_k) | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) = \text{N}(\mathbf{x}(t_k) | \mathbf{m}_k^-, P_k^-), \quad (7)$$

and where the predicted mean and covariance are calculated from

$$\begin{aligned} \mathbf{m}_k^- &= F_{k-1} \mathbf{m}_{k-1}, \\ P_k^- &= F_{k-1} P_{k-1} F_{k-1}^\top + \Sigma_{k-1}. \end{aligned} \quad (8)$$

The measurement \mathbf{y}_k at time t_k is then incorporated to create an updated estimation of the latent state $\mathbf{x}(t_k)$

$$p(\mathbf{x}(t_k) | \mathbf{y}_1, \dots, \mathbf{y}_k) = \text{N}(\mathbf{x}(t_k) | \mathbf{m}_k, P_k) \quad (9)$$

and where the updated mean and covariance are calculated using the update steps

$$\begin{aligned} \mathbf{v}_k &= \mathbf{y}_k - H_k \mathbf{m}_k^-, \\ S_k &= H_k P_k^- H_k^\top + R_k, \\ K_k &= P_k^- H_k^\top S_k^{-1}, \\ \mathbf{m}_k &= \mathbf{m}_k^- + K_k \mathbf{v}_k, \\ P_k &= P_k^- - K_k S_k K_k^\top. \end{aligned} \quad (10)$$

From these equations the terms necessary for the likelihood (Equation 6) can be calculated as a by-product

$$p(\mathbf{y}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) = \text{N}(\mathbf{y}_k | H_k \mathbf{m}_k^-, S_k). \quad (11)$$

We will now specify the three models used, which allow calculation of the terms in Equation 11 and hence the likelihood using Equation 6. For each of the models we use the "LinearGaussianStateSpaceModel" distribution within Tensorflow Probability. This distribution requires the specification of the mean and covariance of both the dynamics and observation models, and the log likelihood is then calculated using the 'log_prob()' method.

3.5 Model 1: glucose and meals model

We model glucose dynamics with a two-dimensional system of SDEs with two variables: x_{GLUC1} and x_{GLUC2} . The second variable x_{GLUC2} acts as an input into the observation model to produce the measured variable y , which corresponds to the continuous glucose monitor (CGM) data. In matrix form, the model is expressed as follows

$$\begin{aligned} d\mathbf{x}(t) &= W\mathbf{x}(t)dt + d\boldsymbol{\beta} \\ \mathbf{x}(t) &= \begin{bmatrix} x_{\text{GLUC1}}(t) \\ x_{\text{GLUC2}}(t) \end{bmatrix}, W = \begin{bmatrix} -A_{11} & -A_{12} \\ A_{21} & -A_{22} \end{bmatrix}, Q = \begin{bmatrix} 0 & 0 \\ 0 & B_{22} \end{bmatrix}, \end{aligned} \quad (12)$$

where the coefficients A_{ij} are constrained to be positive. Given the coefficients in the W matrix, the positive constraints on the coefficients A_{ij} ensures stability of the matrix (i.e. $-A_{11} - A_{22} < 0$ and $A_{11}A_{22} > -A_{12}A_{21}$). There are two principal reasons for modelling the glucose with two variables: 1) the inclusion of a second variable allows for negative feedback via the negative term $-A_{12}$, as negative feedback can arise due to the action of insulin; 2) when meals are consumed, we model this as impacting the first variable x_{GLUC1} . As the first variable x_{GLUC1} acts to increase levels in the second variable x_{GLUC2} (via the parameter A_{21}), this leads to smooth, continuous increases in glucose after food consumption, which is a feature observed in the data. The covariance of the brownian noise term $\boldsymbol{\beta}$ is given by Q .

3.5.1 State-transition model

The state-transition model F_k can be calculated using a 2x2 matrix exponential

$$\begin{aligned} F_k &= \exp(W\Delta t_k), \\ &= e^{s\Delta t_k} \left(\left(\cosh(\sqrt{\gamma}\Delta t_k) - s \frac{\sinh(\sqrt{\gamma}\Delta t_k)}{\sqrt{\gamma}} \right) I + \frac{\sinh(\sqrt{\gamma}\Delta t_k)}{\sqrt{\gamma}} W \right), \end{aligned} \quad (13)$$

where

$$\begin{aligned} s &= (-A_{11} - A_{22})/2, \\ \gamma &= (-A_{11} + A_{22})^2/4 - A_{12}A_{21} = s^2 - \det(W). \end{aligned} \quad (14)$$

If $\gamma < 0$ then

$$\begin{aligned}\cosh \sqrt{\gamma} \Delta t_k &= \cos \sqrt{-\gamma} \Delta t_k, \\ \frac{\sinh \sqrt{\gamma} \Delta t_k}{\sqrt{\gamma}} &= \frac{\sin \sqrt{-\gamma} \Delta t_k}{\sqrt{-\gamma}}.\end{aligned}$$

3.5.2 Covariance of the process noise

The covariance of the process noise Σ_k can be calculated from the steady-state covariance P_∞ as follows

$$\Sigma(\Delta t_k) = P_\infty - F_k(\Delta t_k) P_\infty F_k^\top(\Delta t_k), \quad (15)$$

where P_∞ can be calculated symbolically using the following formula

$$W P_\infty + P_\infty W^\top + Q = \mathbf{0}. \quad (16)$$

3.5.3 Incorporating meal events

The smartphone application provides a list of the recorded ingestion event times $\{t_m\}_{m=1}^M$ for a total of M meals. We model the effect of meal event m as causing an instantaneous increase of size δ_m to variable x_{GLUC1} . Given the delays from eating to reaching the interstitial tissue measured by the CGM, we also include a time lag parameter τ for each participant, which is learnt in addition to the other parameters during model fitting. The response function $r_m(t, t_m, \boldsymbol{\theta})$ to a meal is then given by

$$r_m(t, t_m, \boldsymbol{\theta}) = \begin{cases} \Theta(t - \tau - t_m) \frac{A_{21} \delta_m e^{s(t - \tau - t_m)} \sinh(\sqrt{\gamma}(t - \tau - t_m))}{\sqrt{\gamma}}, & \gamma > 0 \\ \Theta(t - \tau - t_m) \frac{A_{21} \delta_m e^{s(t - \tau - t_m)} \sin(\sqrt{-\gamma}(t - \tau - t_m))}{\sqrt{-\gamma}}, & \gamma < 0 \end{cases}, \quad (17)$$

where the Heaviside function $\Theta(t - \tau - t_m)$ is defined as

$$\Theta(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}. \quad (18)$$

The meal function is then sum of the M individual meals

$$r(t) = \sum_{m=1}^M r_m(t, t_m, \boldsymbol{\theta}). \quad (19)$$

To simplify the interpretation of the perturbation parameters δ_m (and to also make the definition of the prior simpler), we reparameterise the parameters δ_m to correspond to the maximum height of the glucose meal response. The time $t_{r_{\max}}$ of the maximum height of the meal response is

$$t_{r_{\max}} = \begin{cases} \tanh^{-1}(-\sqrt{\gamma}/s)/\sqrt{\gamma}, & \gamma > 0 \\ \tan^{-1}(-\sqrt{-\gamma}/s)/\sqrt{-\gamma}, & \gamma < 0 \end{cases}, \quad (20)$$

and the the maximum height of the glucose meal response is then given as

$$\max(r_m(t, t_m, \boldsymbol{\theta})) = \begin{cases} \frac{A_{21} \delta_m e^{\delta_m (t_{r_{\max}})} \sinh(\sqrt{\gamma} (t_{r_{\max}}))}{\sqrt{\gamma}}, & \gamma > 0 \\ \frac{A_{21} \delta_m e^{\delta_m (t_{r_{\max}})} \sin(\sqrt{-\gamma} (t_{r_{\max}}))}{\sqrt{-\gamma}}, & \gamma < 0 \end{cases} \quad (21)$$

A meal height parameter δ_m^* is then defined as $\delta_m^* = \max(r_m(t, t_m, \boldsymbol{\theta}))$, and the original perturbation parameter δ_m is then found by renormalising.

3.5.4 Circadian dynamics

In addition to the effect of external meal events, we also model underlying circadian rhythms in glucose levels with a sinusoidal function

$$g_{\text{GLUC}}(t) = A_{0,\text{GLUC}} + A_{1,\text{GLUC}}(1 + \cos(\omega t - \phi_{\text{GLUC}}))/2, \quad (22)$$

where $A_{0,\text{GLUC}}$ is the baseline level, $A_{1,\text{GLUC}}$ is the amplitude, ω is the frequency (fixed at $2\pi/24$), and ϕ_{GLUC} is the peak time of the 24-hour function.

3.5.5 Observation matrix

The observation matrix H_k is such that only the second variable is observed

$$H_k = [0, 1]. \quad (23)$$

3.5.6 Observation model

Finally, the observation model has the following form

$$\begin{aligned} \mathbf{y}_k &= H_k \mathbf{x}(t_k) + N(\mathbf{m}_k, R_k) \\ \mathbf{m}_k &= r(t) + g_{\text{GLUC}}(t), R_k = \sigma_{\text{GLUC}}^2. \end{aligned} \quad (24)$$

3.5.7 Parameter description and priors

Parameter	Description	Prior
$A_{0, \text{GLUC}}$	Circadian baseline	$A_{0, \text{GLUC}} \sim \text{half-normal}(5)$
$A_{1, \text{GLUC}}$	Circadian amplitude	$A_{1, \text{GLUC}} \sim \text{half-normal}(1)$
ϕ_{GLUC}	Circadian phase	$\phi_{\text{GLUC}} \sim \text{U}(0, 24)$
A_{11}	$x_{\text{GLUC}1}$ degradation rate	$\log(A_{11}) \sim \text{N}(0, 1)$
A_{12}	Suppression of $x_{\text{GLUC}1}$ by $x_{\text{GLUC}2}$	$\log(A_{12}) \sim \text{N}(0, 1)$
A_{21}	Activation of $x_{\text{GLUC}2}$ by $x_{\text{GLUC}1}$	$\log(A_{21}) \sim \text{N}(0, 1)$
A_{22}	$x_{\text{GLUC}2}$ degradation rate	$\log(A_{22}) \sim \text{N}(0, 1)$
τ	Delay in meal response	$\tau \sim \text{half-normal}(0.5)$
B_{22}	Glucose diffusion noise	$B_{22} \sim \text{half-normal}(0.5)$
δ_m^*	Meal-induced glucose increase	$\delta_m^* \sim \text{half-normal}(5)$

Parameters Model 1 : glucose and meals model.

3.6 Model 2: physical and heart activity model

We model physical and heart activity dynamics with a three-dimensional system of SDEs, where the first variable x_{ACT} represents physical activity, the second variable x_{HR} represents heart rate and the third variable x_{HRV} represents heart rate variability, where we use the inverse of the root mean square of successive differences between normal heartbeats (RMSSD^{-1}). We normalise all three variables by their respective standard deviations before inferring parameters. In matrix form, the model is expressed as follows

$$d\mathbf{x}(t) = W\mathbf{x}(t)dt + d\boldsymbol{\beta},$$

$$\mathbf{x}(t) = \begin{bmatrix} x_{\text{ACT}}(t) \\ x_{\text{HR}}(t) \\ x_{\text{HRV}}(t) \end{bmatrix} W = \begin{bmatrix} -C_{11} & 0 & 0 \\ C_{21} & -C_{22} & 0 \\ C_{31} & 0 & -C_{33} \end{bmatrix} Q = \begin{pmatrix} D_{11} & 0 & 0 \\ 0 & D_{22} & \rho\sqrt{D_{22}D_{33}} \\ 0 & \rho\sqrt{D_{22}D_{33}} & D_{33} \end{pmatrix}, \quad (25)$$

where the coefficients C_{ij} are constrained to be positive. The covariance of the brownian noise term $\boldsymbol{\beta}$ is given by Q . Dynamic changes and fluctuations in HR and HRV might be correlated, where this correlation is mediated by e.g. the autonomic nervous system. In the model, the correlation in the fluctuations between HR and HRV is quantified with the correlation parameter $\rho \in (-1, 1)$.

3.6.1 State-transition model

The state-transition model F_k can be calculated using a 3x3 matrix exponential

$$\begin{aligned}
F_k &= \exp(W \Delta t_k), \\
&= \begin{pmatrix} e^{-C_{11} \Delta t_k} & 0 & 0 \\ -\frac{C_{21} e^{-C_{11} \Delta t_k} - C_{21} e^{-C_{22} \Delta t_k}}{C_{11} - C_{22}} & e^{-C_{22} \Delta t_k} & 0 \\ -\frac{C_{31} e^{-C_{11} \Delta t_k} - C_{31} e^{-C_{33} \Delta t_k}}{C_{11} - C_{33}} & 0 & e^{-C_{33} \Delta t_k} \end{pmatrix}.
\end{aligned} \tag{26}$$

3.6.2 Covariance of the process noise

Similarly to the glucose model, the covariance of the process noise is calculated symbolically using the following formulae

$$\Sigma(\Delta t_k) = P_\infty - F_k(\Delta t_k)P_\infty F_k^\top(\Delta t_k), \tag{27}$$

$$WP_\infty + P_\infty W^\top + Q = \mathbf{0}. \tag{28}$$

3.6.3 Circadian dynamics

We model the underlying circadian rhythms in activity levels with a sinusoidal function

$$g_{\text{ACT}}(t) = C_{0,\text{ACT}} + C_{1,\text{ACT}}(1 + \cos(\omega t - \phi_{\text{ACT}}))/2, \tag{29}$$

where C_0 is the baseline level, C_1 is the amplitude, ω is the frequency (fixed at $2\pi/24$), and ϕ_{ACT} is the peak time of the function. Given that physical activity acts as an input into HR, the circadian function for HR receives an input from physical activity. We therefore model the circadian rhythms for HR using two functions: a term that integrates input from physical activity $g_{\text{HR}}^1(t)$ and an independent function $g_{\text{HR}}^2(t)$. The expression for the term $g_{\text{HR}}^1(t)$ that integrates input from physical activity is then given by

$$\begin{aligned}
\frac{d}{dt}g_{\text{HR}}^1(t) &= C_{11}g_{\text{ACT}}(t) - C_{22}g_{\text{HR}}^1(t), \\
g_{\text{HR}}^1(t) &= \frac{C_0 C_{21}}{C_{22}} + \frac{C_1 C_{21}}{2 C_{22}} - \frac{C_1 C_{21} \omega \sin(\varphi_{\text{ACT}} - \omega t)}{2 (C_{22}^2 + \omega^2)} + \frac{C_1 C_{21} C_{22} \cos(\varphi_{\text{ACT}} - \omega t)}{2 (C_{22}^2 + \omega^2)}.
\end{aligned} \tag{30}$$

We then model the additional circadian component $g_{\text{HR}}^2(t)$ with a separate function

$$g_{\text{HR}}^2(t) = C_{0,\text{HR}} + C_{1,\text{HR}}(1 + \cos(\omega t - \phi_{\text{HR}}))/2. \tag{31}$$

The total circadian function is then a sum of these two components

$$g_{\text{HR}}(t) = g_{\text{HR}}^1(t) + g_{\text{HR}}^2(t). \tag{32}$$

The same framework is then used for the HRV signal, with a term that integrates input from physical activity $g_{\text{HRV}}^1(t)$ and an independent function $g_{\text{HRV}}^2(t)$. The expression for the term $g_{\text{HRV}}^1(t)$ that integrates input from physical activity is then given by

$$\begin{aligned} \frac{d}{dt}g_{\text{HRV}}^1(t) &= C_{11}g_{\text{ACT}}(t) - C_{22}g_{\text{HRV}}^1(t) \\ g_{\text{HRV}}^1(t) &= \frac{C_0 C_{31}}{C_{33}} + \frac{C_1 C_{31}}{2 C_{33}} - \frac{C_1 C_{31} \omega \sin(\varphi_{\text{ACT}} - \omega t)}{2 (C_{33}^2 + \omega^2)} + \frac{C_1 C_{31} C_{33} \cos(\varphi_{\text{ACT}} - \omega t)}{2 (C_{33}^2 + \omega^2)}. \end{aligned} \quad (33)$$

We then model additional HRV circadian component $g_{\text{HRV}}^2(t)$ with a separate function

$$g_{\text{HR}}^2(t) = C_{0,\text{HRV}} + C_{1,\text{HRV}}(1 + \cos(\omega t - \phi_{\text{HRV}}))/2. \quad (34)$$

The total circadian function is then a sum of these two components

$$g_{\text{HRV}}(t) = g_{\text{HRV}}^1(t) + g_{\text{HRV}}^2(t). \quad (35)$$

3.6.4 Observation matrix

The observation matrix H_k is such that all three variables are observed

$$H_k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (36)$$

3.6.5 Observation model

Finally, the observation model has the following form

$$\begin{aligned} \mathbf{y}(t_k) &= H_k \mathbf{x}(t_k) + N(\mathbf{m}_k, R_k), \\ \mathbf{m}_k &= \begin{bmatrix} g_{\text{ACT}}(t) \\ g_{\text{HR}}(t) \\ g_{\text{HRV}}(t) \end{bmatrix}, R_k = \begin{bmatrix} \sigma_{\text{ACT}}^2 & 0 & 0 \\ 0 & \sigma_{\text{HR}}^2 & 0 \\ 0 & 0 & \sigma_{\text{HRV}}^2 \end{bmatrix}. \end{aligned} \quad (37)$$

3.6.6 Parameter description and priors

Parameter	Description	Prior
$C_{0,ACT}, C_{0,HR}, C_{0,HRV}$	Circadian baseline (activity, HR, HRV)	$C_{0,ACT}, C_{0,HR}, C_{0,HRV} \sim \text{half-normal}(1)$
$C_{1,ACT}, C_{1,HR}, C_{1,HRV}$	Circadian amplitude (activity, HR, HRV)	$C_{1,ACT}, C_{1,HR}, C_{1,HRV} \sim \text{half-normal}(1)$
$\phi_{ACT}, \phi_{HR}, \phi_{HRV}$	Circadian phase (activity, HR, HRV)	$\phi_{ACT}, \phi_{HR}, \phi_{HRV} \sim U(0, 24)$
C_{11}, C_{22}, C_{33}	Relaxation rate (activity, HR, HRV)	$C_{11}^{-1}, C_{22}^{-1}, C_{33}^{-1} \sim \text{half-normal}(0.01)$
C_{21}	Activation of x_{HR} by x_{ACT}	$\log(C_{21}) \sim N(0, 5)$
C_{31}	Activation of x_{HRV} by x_{ACT}	$\log(C_{31}) \sim N(0, 5)$
ρ	Correlation in diffusion between x_{HR} and x_{HRV}	$\rho \sim U(-1, 1)$
D_{11}, D_{22}, D_{33}	Diffusion noise (activity, HR, HRV)	$D_{11}, D_{22}, D_{33} \sim \text{half-normal}(5)$
$\sigma_{ACT}, \sigma_{HR}, \sigma_{HRV}$	Measurement noise (activity, HR, HRV)	$\sigma_{ACT}, \sigma_{HR}, \sigma_{HRV} \sim \text{half-normal}(1)$

Parameters Model 2 : physical and heart activity model.

3.7 Model 3: combined model

The final model connects the physical and heart activity signals with glucose dynamics by connecting Model 1 (the glucose and meals model) with Model 2 (physical and heart activity). Three new parameters C_{51}, C_{52} and C_{53} are introduced that describe the effect of physical activity, HR and HRV on glucose levels (x_{GLUC2}), respectively. These three parameters are left unconstrained and can take either positive, negative or zero values.

$$\begin{aligned}
 d\mathbf{x}(t) &= W\mathbf{x}(t)dt + d\boldsymbol{\beta} \\
 \mathbf{x}(t) &= \begin{bmatrix} x_{ACT}(t) \\ x_{HR}(t) \\ x_{HRV}(t) \\ x_{GLUC1}(t) \\ x_{GLUC2}(t) \end{bmatrix} \quad W = \begin{bmatrix} -C_{11} & 0 & 0 & 0 & 0 \\ C_{21} & -C_{22} & 0 & 0 & 0 \\ C_{31} & 0 & -C_{33} & 0 & 0 \\ 0 & 0 & 0 & -A_{11} & -A_{12} \\ C_{51} & C_{52} & C_{53} & A_{21} & -A_{22} \end{bmatrix} \\
 Q &= \begin{bmatrix} D_{11} & 0 & 0 & 0 & 0 \\ 0 & D_{22} & \rho\sqrt{D_{22}D_{33}} & 0 & 0 \\ 0 & \rho\sqrt{D_{22}D_{33}} & D_{33} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & B_{22} \end{bmatrix}.
 \end{aligned} \tag{38}$$

To simplify the model inference problem, the parameters from Model 2 describing the physical and heart activity model in isolation were locked to their posterior mean values. The covariance of the brownian noise

term β is given by Q .

3.7.1 State-transition model

Symbolic calculation of the matrix exponential required for the state-transition model becomes too cumbersome for Model 3, and we therefore use a Taylor series to approximate it [S2].

$$\exp(W\Delta t_k) = \sum_{j=0}^l (W\Delta t_k)^j / j!, \quad (39)$$

along with the scaling property

$$e^{(W\Delta t_k)} = \left(e^{(W\Delta t_k)/m} \right)^m. \quad (40)$$

We use $l = 5$ and $m = 2$.

3.7.2 Covariance of the process noise

For the covariance of the process noise, we use the Matrix Fraction Decomposition approach [S3-S4] to solve

$$\frac{d\Sigma(t)}{dt} = W\Sigma(t) + \Sigma(t)W^\top + Q, \quad \Sigma(0) = \mathbf{0} \quad (41)$$

If we define matrices $C_\Sigma(\Delta t)$ and $D_\Sigma(\Delta t)$ such that $\Sigma(\Delta t) = C_\Sigma(\Delta t)D_\Sigma^{-1}(\Delta t)$ then we can now use matrix fractions to solve

$$\begin{aligned} \begin{pmatrix} C_\Sigma(\Delta t) \\ D_\Sigma(\Delta t) \end{pmatrix} &= \exp \left[\begin{pmatrix} W & Q \\ \mathbf{0} & -W^\top \end{pmatrix} \Delta t \right] \begin{pmatrix} \mathbf{0} \\ \mathbf{I} \end{pmatrix} \\ &= \begin{pmatrix} \exp(W\Delta t) & \Sigma(\Delta t) \exp(W\Delta t)^{-\top} \\ \mathbf{0} & \exp(W\Delta t)^{-\top} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \mathbf{I} \end{pmatrix} \end{aligned} \quad (42)$$

As $D_\Sigma^{-1}(\Delta t) = \exp(W\Delta t)^\top$, there is no need for a matrix inversion.

3.7.3 Observation matrix

The observation matrix H_k is such that physical activity ($x_{\text{ACT}}(t)$), HR ($x_{\text{HR}}(t)$) and HRV ($x_{\text{HRV}}(t)$) are observed along with the glucose levels ($x_{\text{GLUC2}}(t)$).

$$H_k = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (43)$$

3.7.4 Observation model

Finally, the observation model has the following form

$$\mathbf{y}_k = H_k \mathbf{x}(t_k) + N(\mathbf{m}_k, R_k)$$

$$\mathbf{m}_k = \begin{bmatrix} g_{\text{ACT}}(t) \\ g_{\text{HR}}(t) \\ g_{\text{HRV}}(t) \\ g_{\text{GLUC}}(t) \end{bmatrix}, R_k = \begin{bmatrix} \sigma_{\text{ACT}}^2 & 0 & 0 & 0 \\ 0 & \sigma_{\text{HR}}^2 & 0 & 0 \\ 0 & 0 & \sigma_{\text{HRV}}^2 & 0 \\ 0 & 0 & 0 & \sigma_{\text{gluc}}^2 \end{bmatrix} \quad (44)$$

3.7.5 Parameter description and priors

Parameters for the physical and heart activity model (Model 2) are fixed at their posterior mean values, and the parameters for the glucose model (Model 1) are learnt with the same priors. The priors for the 3 new parameters (C_{51}, C_{52}, C_{53}) in Model 3 are then as follows

Parameter	Description	Prior
C_{51}	Activation of $x_{\text{GLUC}2}(t)$ by activity	$C_{51} \sim N(0, 1)$
C_{52}	Activation of $x_{\text{GLUC}2}(t)$ by HR	$C_{52} \sim N(0, 1)$
C_{53}	Activation of $x_{\text{GLUC}2}(t)$ by HRV	$C_{53} \sim N(0, 1)$

Parameters Model 3 : combined model.

3.8 Posterior parameter sampling with MCMC

The parameter posterior distribution was sampled using Hamiltonian Markov Chain Monte Carlo (HMC), which uses the gradients of the posterior to improve the efficiency of the sampling. To initialise the sampler, we found the maximum a posteriori probability (MAP) parameter estimate using the BFGS optimiser 'bfgs_minimize' within TensorFlow Probability. We then used the 'HamiltonianMonteCarlo' function with TensorFlow Probability with 5 leapfrog steps, and we scaled the step size of each variable to approximately match the standard deviation of the posterior distribution. To achieve this, we sampled posterior parameters using two steps. Firstly, we sampled 10000 parameters (with a burn-in of 10,000 samples) using the 'SimpleStepSizeAdaptation' kernel to select the global step size, which adapts the global step size to achieve a target acceptance probability of 0.75 [S5]. We then scaled the step size of each variable according to the standard deviation of this initial posterior distribution. Next, we resampled model parameters from the posterior distribution using 4 different chains with 10,000 samples each (with a burn-in of 10,000 samples), again using 'SimpleStepSizeAdaptation' kernel to globally rescale the step size. The 'SimpleStepSizeAdaptation' kernel was only applied to first 80% of the burn-in samples.

4 References

- S1 Simo Särkkä and Arno Solin (2019). Applied Stochastic Differential Equations. Cambridge University Press.
- S2 Moler C, Van Loan C (2003). Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Rev.*,45: 3-49.
- S3 Simo Särkkä (2006) Recursive Bayesian Inference on Stochastic Differential Equations. Doctoral dissertation, Helsinki University of Technology, Espoo, Finland.
- S4 Axelsson, P. and Gustafsson, F. (2015). Discrete-time solutions to the continuous-time differential Lyapunov equation with applications to Kalman filtering. *IEEE Transactions on Automatic Control*, 60(3), 632-643.
- S5 Andrieu, Christophe, Thoms, Johannes. A tutorial on adaptive MCMC. *Statistics and Computing*, 2008. <https://people.eecs.berkeley.edu/~jordan/sail/readings/andrieu-thoms.pdf>