pp. 1–22 (2022)

Algorithms and methodological challenges in the development and application of quantitative systems pharmacology models: a case study in type 2 diabetes

V. Sokolov*

Abstract — Quantitative systems pharmacology (QSP) is a relatively new modelling discipline, formed within the ever-growing domain of model-informed drug development and actively evolving throughout the last decade. This modelling technique is based on the systems analysis and is used to get a quantitative rather than qualitative understanding of systems dynamics and explore the mechanisms of action of a drug. However, there is no well-defined methodology for the QSP model development, which significantly complicates the practical application of these models. In the current work, we overview the existing mathematical models of antidiabetic therapies and propose a modelling method, which overcomes common limitations and is able to produce a physiologically based mechanistic model describing gliflozin action in type 2 diabetes mellitus. From the practical standpoint, sensitivity analysis preformed in this work helped to reveal subpopulation of patients with better response to gliflozin therapy.

Keywords: Systems analysis, dynamics, ordinary differential equations, quantitative systems pharmacology, mathematical modelling

MSC 2010: 92C32, 92C45, 92D25, 93A10

Until the beginning of the XX century, biology and pharmacology were associated primarily with the empirical knowledge accumulation through the meticulous planning and execution of *in vitro* and *in vivo* experiments. Exploding growth of the computer engineering and the quantities of the available scientific information, coupled with the lack of theoretical basis necessary to transform data into knowledge gave birth to a new type of modelling discipline pharmacometrics, or quantitative pharmacology [7]. In 1969, Lewis Sheiner published the first work describing a computer program for IBM360/50, which was able to predict an optimal daily dose of warfarin, an anti-coagulant, based on the current values of the clotting factors, previous warfarin dose, and physicians guidelines [49]. The model consists of two equations: analytical expression representing pharmacokinetics (PK), or how the body affects a drug and an ordinary differential equation reflecting pharmacodynamics (PD) of clotting factor, or how a drug affects the body.

Through their subsequent works, Sheiner and colleagues formed a modelling approach called ‘population PK/PD modelling’ and developed a computer software

*M&S Decisions, Department of Pharmacometrics, Moscow 125167, Russia. STU Sirius, Research Center of IT and AI, Sochi 354349, Russia. E-mail: victor.sokolov@msdecisions.ru

called NONMEM able to implement nonlinear mixed effects modelling approach for the quantitative analysis of PK and PD data [51]. Nowadays, this modelling technique is an integral part of the development of new medicines [42]. In Sheiners 'Learn and confirm' paradigm, drug development process should be viewed as two sequential cycles of induction (learning) and deduction (confirming), drawing inference from the former primarily on the basis of statistical and mechanistic models [50]. Application of such models to the inverse (i.e., parameter estimation) and direct (e.g., clinical trial simulations) modelling problems significantly accelerated drug development process and transformed regulatory decision-making [30, 34].

Application of the systems analysis theory to the biological problems provided the basis for the quantification of emergent physiological properties through the mechanistic description of the interplay and dynamical behavior of system components. Among these methods, quantitative systems pharmacology (QSP) stands out as a versatile tool, capable of answering various essential questions throughout the drug development process and beyond, such as identification of optimal biological targets, validation of a mechanism of action, choice of the optimal population to treat, competitor benchmarking. The primary goal of a QSP model is to quantitatively and mechanistically characterize biological, toxicological, and disease processes in response to therapeutic modulation [9]. Therefore, QSP model can be defined as a mathematical system, which integrates pharmacological preclinical and clinical data to predict dynamical and spatial pathophysiological response. The spectrum of modelling approaches considered as QSP is not well-determined, and sometimes includes even computational fluid dynamics models, agent-based modelling and machine learning [38]. Nevertheless, the core of the method may be best represented by the mechanism-based pathway and signal-transduction models and mechanistic PK/PD systems of various scales, based on the deterministic ordinary differential equations (ODE) or delayed differential equations (DDE).

Obligatory robustness of regulatory-level decisions creates multiple challenges for QSP model development and application. Among them, the details on expert knowledge and data used for model development, differences in model scope and complexity, lack of acceptance criteria and common expectations of model credibility, and discrepancies in solving of the direct problem for virtual patient populations can be considered as the most common and essential [2]. In current research, we propose a methodology able to overcome part of these limitations and illustrate it with an integrative QSP model of type 2 diabetes mellitus (T2DM).

Diabetes mellitus (DM) is a group of diseases characterized by an abnormally high glucose concentration in blood, affecting hundreds of millions of people worldwide [1]. In healthy condition, average glucose concentration in blood is maintained in a narrow interval through a complex neuro-humoral regulatory system [46]. In the most commonly occurring T2DM, tissue resistance prevents insulin from stimulating glucose utilization, resulting in an increase in insulin and glucose plasma concentration, at least during the early phases of the disease [12].

A lot of mathematical models were developed to explore various aspects of T2DM [26]. Among them, the Bergman minimal model, published in 1979, stands

out as one of the first examples of *in silico* description of glucose-insulin homeostasis: the cornerstone of the pathophysiology of DM [8]. The original model contained 3 ordinary differential equations and 4 parameters describing the dynamical changes of insulin in blood, insulin in an effect compartment, and glucose in blood after intravenous glucose tolerance test in dogs. The Bergman minimal model was subsequently updated to describe insulin-sensitivity in humans and used as a basis for multiple model extensions and applications. One of the initial model limitation caused by a necessity to calibrate parameters sequentially using forcing functions to describe insulin and glucose time series was later addressed by De Gaetano and Arino [11]. Jauslin and Silber subsequently updated this mathematical model to describe glucose-insulin crosstalk in healthy subjects and patients with T2DM after intravenous, oral glucose tolerance test, and food intake [22–24, 52, 53].

Life-threatening complications of T2DM occur after continuous exposure to the high glucose levels in blood, which are in turn reflected in the proportion of glycated haemoglobin (HbA1c). As such, a few mathematical models were developed to describe the kinetics of haemoglobin glycation [6, 19, 21, 29, 31, 35, 36, 39, 55]. One approach considers the formation of HbA1c as a function of red blood cells life span and blood glucose concentration based on the data from different populations of healthy subjects and DM patients [29].

Many mathematical models were also created to support the development of anti-T2DM therapies. In particular, the successful release of the gliflozins to the market was accompanied by a set of modelling works based on empirical or semi-empirical PK/PD models [4, 13, 44, 45, 56]. Gliflozins (or SGLT2 inhibitors, SGLT2i) decrease glucose concentration in blood by inhibiting glucose reabsorption in kidneys and thereby causing the excess of glucose to be excreted with urine. Therefore, physiologically based models of gliflozin effect contain mathematical description of kidney lumen, processes of renal filtration, reabsorption, and urine secretion. Gliflozins are also one of the rare examples when positive effects of the treatment are being discovered even after compounds release to the market: treatment with SGLT2i reduced the risks of heart failure; the nature of this phenomenon was also explained via mathematical model [18]. Finally, further *in silico* research was substantiated by the potential application of the gliflozins as an add-on therapy to insulin in type 1 DM [25, 43].

A mathematical platform of renal glucose reabsorption, insulin-glucose homeostasis, and haemoglobin glycosylation in T2DM highlighted in this research is an example of an integrative QSP model designed to capture various aspects of the disease. The details on model development are described elsewhere [48]. In this study, we perform a sensitivity analysis using the platform to identify subgroup of patients most responsive to the treatment with gliflozins.

1. Computational methods of QSP model development

QSP is a relatively new modelling technique. The term QSP was first proposed by the National Institute of Health, USA in 2011 white paper [58]. Since then, the

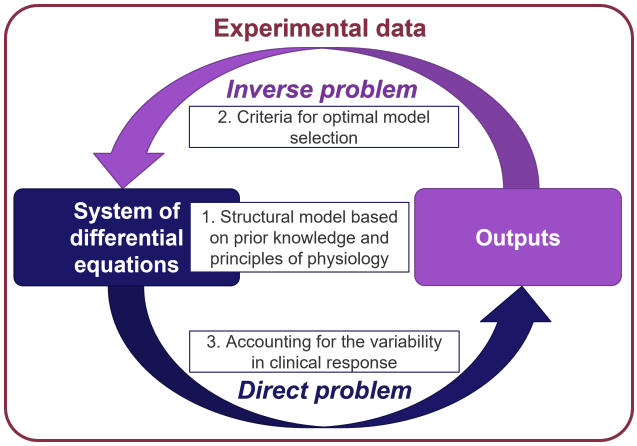

Figure 1. Generalized scheme of model development and associated limitations in QSP.

methodology became widespread within the pharmaceutical industry and academic researchers. However, despite its frequent use, the methods and algorithms, workflow, software tools, and criteria for quality assessment of QSP models are still loosely defined, which in turn limits further application of QSP modelling for regulatory decision making [2, 15, 38]. Three categories of common limitations were derived, related to the data handling and definition of a structural model, identifiability of parameters, and forward simulations (see Fig. 1) and a more generalized methodology able to overcome these limitations was proposed.

1.1. Defining the structural and statistical models

In a way, QSP modelling is an embodiment of Bayesian principle, as it integrates all prior knowledge on a physiological system to make inference on new potential scenarios. In case of QSP modelling, prior knowledge first of all includes all up-to-date information on key pathophysiological processes within a disease area, relevant biological entities such as molecules and cells, their spatial distribution and interactions, and all possible quantitative data related to these processes. The quantitative data for model development should be collected through a process of systematic literature review, following the same guidelines created for the meta-analyses (PRISMA guidelines): dates, search queries, exclusion/inclusion criteria, number of papers, etc., should be indicated and validated [40]. Trial design, population characteristics and time series of different measurements should be digitized into a standardized database and processed into the datasets, acceptable by a software, through reproducible scripts. Following these actions, a modeler can define the necessary set of model variables within the system of interest, relate observed measurements with those variables, evaluate the dynamics and population distributions, correlation between the measurements, between-study and between-subject variability, and infer possible functional relationships and feedbacks between the model components [20].

In contrast to the phenomenological approach typically used in the population PK/PD modelling, the structure of a QSP model should be based on the principles of (1) enzymological kinetics to reflect nonlinearity and regulations specific to biological reactions; (2) steady-state conditions corresponding to the systems homeostasis; (3) multicompartmental structure mimicking physiological distribution of biomarkers and drug exposure. Typical actualization of the first principle is the application of Michaelis–Menten equation when describing the interactions of a substrate, enzyme, and a drug (competitive inhibitor) within the model [10]:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{k_{\mathrm{cat}} E_0 S(t)}{Km(1 + I(t)/Ki) + S(t)} \tag{1.1}$$

where k_{cat} is the catalytic constant, Km is the Michaelis constant, Ki is the inhibition constant, P is the product of the enzymatic reaction, E_0 is the concentration of an enzyme, $S(t)$ is the concentration of a substrate, $I(t)$ is the concentration of an inhibitor (e.g., a drug).

Second principle is illustrated by deriving initial conditions for the variables through model parameters. If $S(t)$ in equation (1.1) is defined by the following ODE in the absence of the enzyme:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = k_{\mathrm{in}} - k_{\mathrm{out}} S(t) \tag{1.2}$$

where k_{in} and k_{out} are synthesis and elimination rate constants, respectively, then initial condition for $S(t)$ should be defined by

$$S(0) = \frac{k_{\mathrm{in}}}{k_{\mathrm{out}}}. \tag{1.3}$$

Finally, it should be noted that the concentrations of biological entities in equations (1.1) and (1.2) can be attributed towards a specific physiological compartment, e.g., kidney tubules, thereby illustrating the third proposed principle of QSP model development. Choosing the model structure based on these principles allows a researcher to move from the apparent macro-parameters to physiologically based micro-parameters measurable in a clinical or preclinical studies. For example, concentration of an enzyme and a substrate, Km, Ki, k_{cat}, and k_{out} parameters can be estimated in the *in vivo* or *in vitro* experiments.

The remaining unknown parameters in the system are identified using the maximum likelihood approach against the observed data, usually available in a form of a time series of measurements $Y(t)$. As such, the parameter estimation procedure can be defined as a search for a vector of parameters Θ for which the predicted outcome $\widehat{Y}(t,\Theta)$ is close to the given experimental data for the matching time points under the following conditions: the residuals between the observed and model-predicted values are normally distributed, and the errors between the observations and between

the variables are independent. Then, the objective function can be described as:

$$L(\Theta) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{1}{2\sigma^2}(Y_j(t) - \widehat{Y}_j(t,\Theta))^2 \Big)$$

$$-2\log(L(\Theta)) = n\log(2\pi) + \sum_{j=1}^{n} \left(\log(\sigma^2) + \frac{(Y_j(t) - \widehat{Y}_j(t,\Theta))^2}{\sigma^2} \right)$$

where j is the jth measurement, n is the total number of measurements, σ is the standard deviation of the residual error, Y_j is the observed value, $\widehat{Y}_j(t,\Theta)$ is the predicted value, Θ is the vector of model parameters, t is the time, and discrimination between the models can be based on the Akaike information criterion (AIC) [16]:

$$AIC = -2\log(L(\Theta)) + 2P$$

where P is the number of unknown parameters in the model.

The variance of the observational error in PK and PD data is not necessarily constant over time, and often increases in proportion to the variable value. Therefore, a residual error model g must be introduced and tested, such that $PRED_j = \widehat{Y}_j(t,\Theta) + g\,\varepsilon_j$, where $\varepsilon \sim N(0,\sigma^2)$, $g = 1$ and $g = \widehat{Y}_j$ for constant and proportional error models, respectively. Another important feature of the QSP modelling is the ability to integrate heterogeneous data, both subject-level (i.e., time series per individual) and aggregated (i.e., mean, or median measurements), from different experiments into a single quantitative system. It substantiates the need to use hierarchical modelling approach through the implementation of random effects, considering physiological constraints of the parameters *via* their transformation:

$$\widehat{\vartheta}_i = \widehat{\vartheta}_{TV} + \eta_i$$

for unconstrained parameters (e.g., power function parameters),

$$\log(\widehat{\vartheta}_i) = \log(\widehat{\vartheta}_{TV}) + \eta_i$$

for the majority of physiologically based parameters,

$$\log\left(\frac{\widehat{\vartheta}_i}{1-\widehat{\vartheta}_i}\right) = \log\left(\frac{\widehat{\vartheta}_{TV}}{1-\widehat{\vartheta}_{TV}}\right) + \eta_i$$

for the parameters constrained between 0 and 1 (e.g., bioavailability or fraction unbound), where $\widehat{\vartheta}_i$ is the estimated individual parameter value, $\widehat{\vartheta}_{TV}$ is the estimated typical parameter value, η_i is the ith random effect, $\eta \sim N(0,\omega^2)$.

Likelihood function cannot be computed in closed form for a non-linear mixed effects modelling (NLME) problem. Thus, QSP model calibration based on individual data requires special algorithms, among which first-order conditional estimation (FOCE) and stochastic approximation expectation maximization (SAEM) methods, implemented in the NONMEM and the Monolix software, are the most commonly used [5, 28].

1.2. Practical identifiability and applicability

Arranging the models by AIC is the first among multiple criteria in QSP model assessment. A model can provide lower AIC score while suffering from the large degree of uncertainty in estimated parameters, which deems the model unusable for the decision-making in drug development, especially if high uncertainty is associated with the drug effect parameters. Thus, at least one of the three techniques to evaluate confidence intervals (CI) for the model parameters must be applied to consider a QSP model usable: variance–covariance matrix, profile likelihood, or bootstrap (parametric or non-parametric) [3, 14, 54].

Variance-covariance matrix can be derived from the Fisher information matrix, in turn calculated either by stochastic approximation using a Markov chain Monte Carlo algorithm, or linearization using a Taylor expansion:

$$I(\widehat{\Theta}) = -\frac{\mathrm{d}^2}{\mathrm{d}\Theta^2}\log(L(\widehat{\Theta})), \quad C(\widehat{\Theta}) = I(\widehat{\Theta})^{-1}, \quad \overline{C}(\widehat{\Theta}) = J^T C(\widehat{\Theta}) J$$

$$\mathrm{SE}(\widehat{\vartheta}_k) = \sqrt{\overline{C}_{k,k}(\widehat{\Theta})}, \quad \mathrm{RSE}(\widehat{\vartheta}_k) = \frac{\mathrm{SE}(\widehat{\vartheta}_k)}{\widehat{\vartheta}_k}\cdot 100\%, \quad \mathrm{CI}(\widehat{\vartheta}_k) = \widehat{\vartheta}_k \pm t_{\alpha,n_d-n_p}\,\mathrm{SE}(\widehat{\vartheta}_k)$$

where $\widehat{\Theta}$ is the vector of parameter estimates, J is the Jacobian, $\overline{C}_{k,k}$ are the diagonal elements of the variance-covariance matrix, n_d is the number of data points, n_p is the number of parameters, t_{α,n_d-n_p} is the percentage point of the Students t-distribution with $n_d - n_p$ degrees of freedom.

Identifiability is assessed through the measure of correlation between the estimated parameters:

$$\mathrm{corr}(\widehat{\vartheta}_{k_1}, \widehat{\vartheta}_{k_2}) = \frac{\overline{C}_{k_1,k_2}}{\mathrm{SE}(\widehat{\vartheta}_{k_1})\cdot\mathrm{SE}(\widehat{\vartheta}_{k_2})}$$

where $\overline{C}_{k_1,k_2}$ is the covariance between parameters $\widehat{\vartheta}_{k_1}$ and $\widehat{\vartheta}_{k_2}$.

Non-parametric bootstrap is based on resampling of the original dataset used for the model calibration into smaller subsets with uniform probability of drawing a random measurement. Parametric bootstrap creates new data samples by perturbing the original data, i.e., $Y_j(t) = Y_j(t) + \sigma N(0,1)$. In both cases, confidence intervals for the parameter estimates are calculated using the following equation:

$$\mathrm{CI}(\widehat{\vartheta}_k) = \mu(\widehat{\vartheta}_k) \pm z_\alpha\,\mathrm{SD}(\widehat{\vartheta}_k)$$

where $\mu(\widehat{\vartheta}_k)$ and $\mathrm{SD}(\widehat{\vartheta}_k)$ are mean and standard deviation of the estimates of $\widehat{\vartheta}_k$, respectively, and z_α is the α percentile of a normal deviate.

Confidence intervals based on the likelihood profiling correspond to the interval $[\vartheta_k^{\min}, \vartheta_k^{\max}]$ of maximal width containing $\widehat{\vartheta}_k$ such that:

$$|\log(L(\dot{\Theta})) - \log(L(\Theta))| \leqslant \frac{1}{2}\chi^2_{1,0.95}$$

where $L(\dot{\Theta})$ is the likelihood when $\widehat{\vartheta}_k$ is fixed within the set of pre-defined values, and $\chi^2_{1,0.95} = 3.84$.

Both bootstrapping and likelihood profiling methods might not be optimal for computations in the framework of complex QSP models; instead, a multi-start parameter estimation procedure to confirm the uniqueness of a found solution can be used.

The presence of highly correlated parameters ($|\mathrm{corr}(\widehat{\vartheta}_{k_1}, \widehat{\vartheta}_{k_2})| \geqslant 0.5$), parameters with RSE $> 50\%$ or 95% CI including zero, and model convergence to different local minima substantially undermines its predictive power and usefulness. Historical data and physiological estimates of the parameter values are therefore vital for avoiding overparametrization of a QSP model, further substantiating the need for thorough systematic review of the published information. Finding more data, fixing parameter values or the distribution of their random effects allow for better identifiability of the remaining unknown parameters, while the contribution of the fixed parameters to the model outputs can be subsequently evaluated through a local or global sensitivity analysis.

Finally, the model quality is assessed by the number of goodness of fit plots, designed to reflect the quality of data description by the model, and include the visualization of observed versus predicted values, weighted residuals over time and over predicted values, distribution of the residuals, visual predictive checks, etc. (see [37]). Even with the best AIC and identifiable parameters, a QSP model may not be able to reflect the trends in the observed data properly. Systemic bias in predictions, skewed distributions, heteroscedasticity in residuals require revisions of the structural and statistical models. Until all three requirements are met, a model cannot be used for model-informed drug development.

1.3. Solving the direct problem

One of the distinct characteristics of the QSP, differentiating it from the other methods of the quantitative pharmacology, is the potential for model extrapolation in the predictions of various treatment scenarios. Therefore, validation of a QSP model by comparing model predictions with the data, previously not used in solving the inverse problem (i.e., external validation) is an advisable step in model development. Furthermore, a taxonomy of model simulations should be developed, to achieve the correct interpretation of the *in silico* scenarios (see Fig. 2).

Ground level of the proposed taxonomy is a model simulation based only on the typical values of the parameters. This simulation corresponds to a population trend in a single study or a cohort and is used in the model evaluation tasks such as sensitivity analyses. Next, by creating a sample of population parameter values from the variance-covariance matrix (see Section 1.2), a modeler can compute the confidence interval around the trend, which is the equivalent of comparing population curves between multiple studies or cohorts. On the other hand, if random effects were introduced, they can be used to explore data heterogeneity, e.g., differentiate subject-level variability from noise, and, combined with the measurements of un-

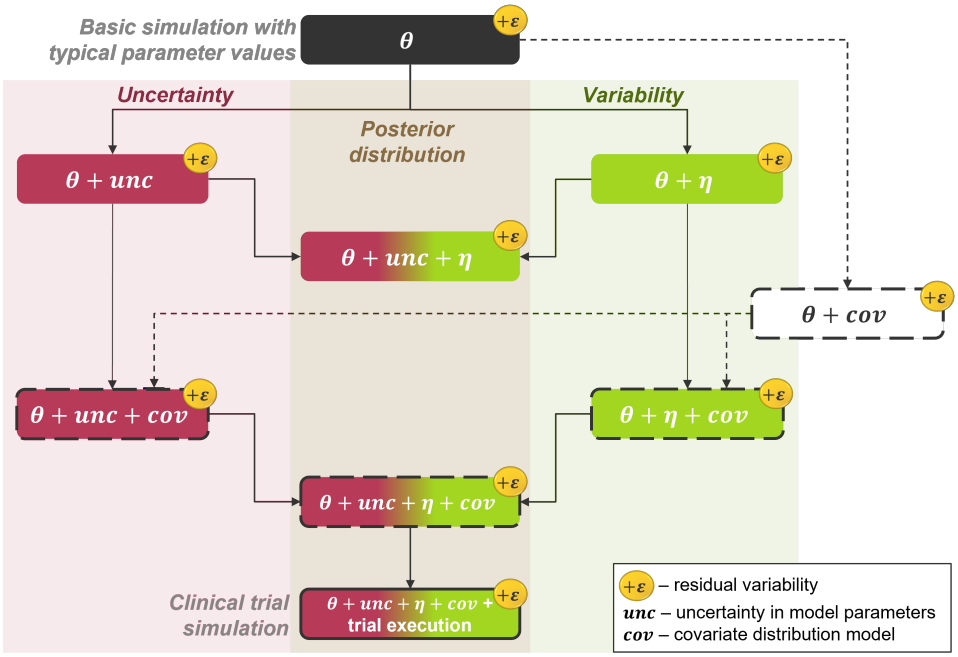

Figure 2. The taxonomy of model-based simulations.

Dapagliflozin PK submodel

Variable	Description	Dimension
$Dapa_d$	Dapagliflozin in dosing compartment	mmol
$Dapa_{pl}$	Dapagliflozin in plasma	mmol
$Dapa_p$	Dapagliflozin in peripheral compartment	mmol

Renal submodel

Variable	Description	Dimension
Glu_{tum1}	Glucose in proximal convoluted tubules	mmol
Glu_{tum2}	Glucose in straight convoluted tubules	mmol
Glu_{bl}	Glucose in bladder	mmol
Glu_{urine}	Glucose in urine	mmol
$Dapa_{tum1}$	Dapagliflozin in proximal convoluted tubules	mmol
$Dapa_{tum2}$	Dapagliflozin in proximal straight tubules	mmol
$Dapa_{bl}$	Dapagliflozin in bladder	mmol
$Dapa_{urine}$	Dapagliflozin in urine	mmol

Glucose homeostasis submodel

Variable	Description	Dimension
$Glucose_d$	Glucose in dosing compartment	mmol
$Glucose_{tr}$	Glucose in the gut	mmol
$Glucose_{pl}$	Glucose in plasma	mmol
$Glucose_p$	Glucose in peripheral compartment	mmol
$Glucose_e$	Glucose in effect compartment	mmol/L
Ins_{pl}	Insulin in plasma	mmol
Ins_e	Insulin in effect compartment	mmol/L

Hemoglobin glycosylation submodel

Variable	Description	Dimension
Hb_n	Hemoglobin $1 \le n \le 12$	-
$HbA1c_n$	Glycosylated hemoglobin $1 \le n \le 12$	-

Figure 3. Scheme and variables of the integrative platform of T2DM.

certainty, derive informative posterior distributions through the model simulations.

Described scenarios can be further updated by introducing covariates: population or trial design characteristics, such as age, body weight, etc., affecting typical parameter values within the model. Proper covariate model can explain part of the variability within the data and allows to perform predictions while simultaneously controlling certain aspects of a virtual population. Together with the elements of the

trial execution, uncertainty, and variability, such predictions can be considered the *in silico* clinical trial simulations.

Finally, all types of the predictions may or may not include residual error, depending on the needs to observe full theoretical variability for the system.

Taken together, the proposed methodology addresses common limitations of the QSP model development by formulating key principles underlying the choice of a structural model, providing a list of criteria for the assessment of model quality, and creating a taxonomy of forward simulations.

2. Mechanistic integrative model of type 2 diabetes

Following prior knowledge existing in the field and previously published modelling analyses, the methodological principles shown above were applied to build an integrative systems model able to link together different aspects of T2DM pathology and various mechanisms of drug action. As a result, a QSP model of gliflozin PK, renal filtration, insulin-glucose homeostasis, and haemoglobin glycation was developed (see Fig. 3) [48]. The model is comprised of 42 ODE and 48 parameters with 20 random effects. Schematic representation of the model structure and variables is provided in Fig. 3. Table 1 contains the data on model parameters, where 32 parameters were fixed based on the previously reported values and 16 parameters were estimated using Nelder–Mean simplex algorithm or SAEM procedure, implemented in Monolix software (version 2019R1) [28]. Solving of the ODE system was performed using backward differentiation formula. Processing of the modelling outputs, forward simulations and other programming activities were performed in R (version 3.5.1).

Table 1. Parameters of the integrative platform of T2DM. Coefficient of variation (CV%) is calculated based on the value of ω, using the following formula for log-transformed parameters: $\mathrm{CV}\% = \sqrt{\exp(\omega^2) - 1} \cdot 100$.

Parameter	Description	Parameter value (CV%)	RSE, %	Dimension	Reference
fup_{Dapa}	Free fraction of dapagliflozin in plasma	0.086	—	—	[57]
MW_{Dapa}	Dapagliflozin molecular weight	408.87	—	g/mol	[57]
k_a^{Dapa}	Dapagliflozin absorption constant	2.39 (156)	—	1/h	[33]
CL_{Dapa}	Dapagliflozin clearance	16.8 (27.5)	4.81	L/h	estimated
Vd_{Dapa}	Dapagliflozin volume of distribution	68.5	5.96	L	estimated
$\beta_{Vd_{Dapa}}$	Body weight effect on Vd_{Dapa}	0.0101	39.8	—	estimated
Vp_{Dapa}	Dapagliflozin volume of peripheral compartment	149	13.7	L	estimated
Q_{Dapa}	Dapagliflozin intercompartmetal clearance	8.42	5.5	L/h	estimated

Continued on next page

Parameter	Description	Parameter value (CV%)	RSE, %	Dimension	Reference
A_{lag}	Dapagliflozin lag time	0.447	1.99	h	estimated
GFR	Glomerular filtration rate	Individual per subject	—	L/h	covariate
BW	Body weight	Individual per subject	—	kg	covariate
		Renal submodel			
V_{lumen1}	Proxima convoluted tubules volume	0.045	—	L	[57]
V_{lumen2}	Proximal straight tubules volume	0.019	—	L	[57]
$V_{bladder}$	Bladder volume	0.2	—	L	[57]
Q_{lumen}	Flux in proximal tubules	2.7	—	L/h	[57]
$Q_{bladder}$	Flux in distal tubules	0.72	—	L/h	[57]
Q_{urine}	Urine excretion	0.055	—	L/h	[57]
Km_{SGLT2}	Michaelis constant for SGLT2	4	—	mmol/L	[57]
Km_{SGLT1}	Michaelis constant for SGLT1	0.5	—	mmol/L	[57]
$Vmax_{tot}^{t2d}$	Maximum SGLT reabsorption capacity	140	—	mmol/L	[57]
Ki_{SGLT1}^{Dapa}	SGLT1 inhibition constant	36.35	—	nmol/L	[57]
Ki_{SGLT2}^{Dapa}	SGLT2 inhibition constant	0.031	—	nmol/L	[57]
$Vmax_{SGLT2}$	Maximum SGLT2 capacity	111(11.6)	1.88	mmol/L	estimated
		Glucose homeostasis submodel			
N	Power function in the modulation function	4	—	—	[47]
Vd_{Glu}	Glucose volume of distribution	9.33 (8.89)	—	L	[24]
Q_{Glu}	Glucose intercompartmental clearance	26.5 (82.9)	—	L	[24]
Vd_{ins}	Insulin volume of distribution	6.09 (16.9)	—	L	[24]
CL_{ins}	Insulin clearance	73.2 (8.42)	—	L/h	[24]
Vp_{Glu}	Glucose volume of peripheral compartment	8.56 (8.93)	—	L	[24]
ke_{Glu}	Delay constant for the glucose effect	0.738 (28.6)	—	1/h	[23]
ke_{ins}	Delay constant for the insulin effect	0.464 (12.2)	—	1/h	[23]
$Fbio_{Glu}$	Glucose bioavailability (oral glucose tolerance test)	0.8	—	—	[41]
$Fbio_{Glu}$ (meal)	Glucose bioavailability (meal)	0.78	—	—	[22]
CL_{Glu}	Insulin-independent glucose clearance	1.72 (35.9)	—	L/h	[24]
$IPRG$	Insulin synthesis	1.42 (12.2)	—	—	[24]

Continued on next page

Parameter	Description	Parameter value (CV%)	RSE, %	Dimension	Reference
GSS	Glucose steady state	6.53 (19.5)	2.91	mmol/L	estimated
ISS	Insulin steady state	9.14 (44.3)	6.38	mU/L	estimated
CL_{Glu}^{Ins}	Insulin-dependent glucose clearance	0.423 (39.6)	6.05	L/(h·mU/L)	estimated
S_{incr}	Incretin effect on insulin secreation	0.01 (27.6)	5.03	—	estimated
ka_{Glu}	Glucose absorption from the gut (oral glucose tolerance test)	1.68 (18.8)	3.28	1/h	estimated
ka_{Glu} (meal)	Glucose absorption from the gut (meal)	0.564 (28.6)	4.57	1/h	estimated
MA	Amplitude of the glucose modulation function	74.6	1.27	—	estimated
MT	Peak time of the glucose modulation	8 h 46 min	0.369	clock time	estimated
MW	Width of the glucose modulation	1.79	1.26	h	estimated
	Haemoglobin glycosylation submodel				
LS	Red blood cells life span	91.7 (0.7)	—	d	[29]
LSP	Red blood cell precursors life span	8.2 (1.32)	—	d	[29]
Kin	Rate of red blood cells synthesis	1	—	1/d	[29]
KG	Glycosylation rate	$8.37{\cdot}10^{-6}$	—	1/(d·mg/dL)	[29]
γ	Glucose feedback on red blood cells life span	−0.38	—	—	[29]
NC	Number of transit compartments	12	—	—	[29]

Dapagliflozin PK in plasma is represented by a two-compartment PK model with first-order absorption and linear elimination (2.1). The dose is administered to the $Dapa_d$ compartment with a lag parameter A_{lag} [h], so that the actual time of dosing $\widehat{T_d} = T_d + A_{lag}$, where T_d [h] is the nominal time of dosing:

$$\begin{aligned}
&\frac{\mathrm{d}Dapa_d}{\mathrm{d}t} = -k_a^{Dapa}\cdot Dapa_d(t), \qquad \frac{\mathrm{d}Dapa_p}{\mathrm{d}t} = Q_{Dapa}\cdot\left(\frac{Dapa_{pl}(t)}{\widehat{Vd_{Dapa}}} - \frac{Dapa_p(t)}{Vp_{Dapa}}\right)\\
&\frac{\mathrm{d}Dapa_{pl}}{\mathrm{d}t} = k_a^{Dapa}\cdot Dapa_d(t) - CL_{Dapa}\cdot\frac{Dapa_{pl}(t)}{\widehat{Vd_{Dapa}}} - GFR\cdot fup_{Dapa}\cdot\frac{Dapa_{pl}(t)}{\widehat{Vd_{Dapa}}}\\
&\qquad\qquad - Q_{Dapa}\cdot\left(\frac{Dapa_{pl}(t)}{\widehat{Vd_{Dapa}}} - \frac{Dapa_p(t)}{Vp_{Dapa}}\right)
\end{aligned} \tag{2.1}$$

where $\widehat{Vd_{Dapa}} = Vd_{Dapa}\cdot\exp\left(\beta_{Vd_{Dapa}}(BW - 83.65)\right)$.

Renal filtration module serves as a basis for the platform. Four compartments of the module represent two segments of the proximal tubules, bladder, and urine,

with the first-order rate of glucose and dapagliflozin appearance in each of them. Glucose reabsorption from kidney tubules back to blood is carried out by two transport proteins: SGLT2 and SGLT1 in the first and the second segments of kidney lumen, respectively (see [17]). Thus, following the principles of enzyme kinetics and considering the mechanisms of action of gliflozins, glucose reabsorption rate is described by a rather complex equations based on Michaelis–Menten kinetics with competitive inhibition from dapagliflozin concentration in respective compartments:

$$\begin{aligned}
\frac{\mathrm{d}Glu_{lum1}}{\mathrm{d}t} &= GFR\cdot\frac{Glucose_{pl}(t)}{\widehat{Vd_{Glu}}} - Q_{lumen}\cdot\frac{Glu_{lum1}(t)}{V_{lumen1}} - V_{RGR1}\\
\frac{\mathrm{d}Glu_{lum2}}{\mathrm{d}t} &= Q_{lumen}\cdot\frac{Glu_{lum1}(t)}{V_{lumen1}} - Q_{bladder}\cdot\frac{Glu_{lum2}(t)}{V_{lumen2}} - V_{RGR2}\\
\frac{\mathrm{d}Glu_{bl}}{\mathrm{d}t} &= Q_{bladder}\cdot\frac{Glu_{lum2}(t)}{V_{lumen2}} - Q_{urine}\cdot\frac{Glu_{bl}(t)}{V_{bladder}}\\
\frac{\mathrm{d}Glu_{urine}}{\mathrm{d}t} &= Q_{urine}\cdot\frac{Glu_{bl}(t)}{V_{bladder}}
\end{aligned}$$

$$\begin{aligned}
\frac{\mathrm{d}Dapa_{lum1}}{\mathrm{d}t} &= GFR\cdot fup_{Dapa}\cdot\frac{Dapa_{pl}(t)}{V_{plasma}} - Q_{lumen}\cdot\frac{Dapa_{lum1}(t)}{V_{lumen1}}\\
\frac{\mathrm{d}Dapa_{lum2}}{\mathrm{d}t} &= Q_{lumen}\cdot\frac{Dapa_{lum1}(t)}{V_{lumen1}} - Q_{bladder}\cdot\frac{Dapa_{lum2}(t)}{V_{lumen2}}\\
\frac{\mathrm{d}Dapa_{bl}}{\mathrm{d}t} &= Q_{bladder}\cdot\frac{Dapa_{lum2}(t)}{V_{lumen2}} - Q_{urine}\cdot\frac{Dapa_{bl}(t)}{V_{bladder}}\\
\frac{\mathrm{d}Dapa_{urine}}{\mathrm{d}t} &= Q_{urine}\cdot\frac{Dapa_{bl}(t)}{V_{bladder}}
\end{aligned}$$

where

$$\begin{aligned}
\widehat{Vd_{Glu}} &= Vd_{Glu}\cdot\frac{BW}{70}\\
V_{RGR1} &= \frac{Vmax_{SGLT_2}\cdot\dfrac{Glu_{lum1}(t)}{V_{lumen1}}}{Km_{SGLT_2}\cdot\left(1+\dfrac{\frac{Dapa_{lum1(t)}}{V_{lumen1}}}{Ki^{Dapa}_{SGLT_2}}\right)+\dfrac{Glu_{lum1}(t)}{V_{lumen1}}}\\
V_{RGR2} &= \frac{Vmax_{SGLT_1}\cdot\dfrac{Glu_{lum2}(t)}{V_{lumen2}}}{Km_{SGLT_1}\cdot\left(1+\dfrac{\frac{Dapa_{lum2(t)}}{V_{lumen2}}}{Ki^{Dapa}_{SGLT_1}}\right)+\dfrac{Glu_{lum2}(t)}{V_{lumen2}}}.
\end{aligned}$$

Overall structure of the glucose-insulin homeostasis module was based on the published IGI model [22]. Equation (2.2) is central for this block, as it describes the

change in plasma glucose over time, taking into account glucose absorption from the gut, gluconeogenesis, insulin-dependent and insulin-independent glucose clearance, renal filtration, and reabsorption. Glucose consumption with food and absorption in the gut is described by two compartments, to capture the delay in glucose appearance after food intake and implement the effect of the incretin on insulin production. Insulin-dependent clearance depends on the insulin concentration in the effect compartment. Insulin secretion is modulated by the glucose in effect compartment and the effect of incretins, mediated by the glucose in the gut:

$$
\begin{aligned}
\frac{\mathrm{d}Glucose_d}{\mathrm{d}t} &= -k_a^{Glu}\cdot Glucose_d(t)\\
\frac{\mathrm{d}Glucose_{tr}}{\mathrm{d}t} &= k_a^{Glu}\cdot Glucose_d(t) - k_a^{Glu}\cdot Glucose_{tr}(t)\\
\frac{\mathrm{d}Glucose_{pl}}{\mathrm{d}t} &= k_a^{Glu}\cdot Glucose_{tr}(t) + GPRO\cdot(1+M) - CL_{Glu}\cdot\frac{Glucose_{pl}(t)}{\widehat{Vd_{Glu}}}\\
&\quad - CL_{Glu}^{Ins}\cdot Ins_e(t)\cdot\frac{Glucose_{pl}(t)}{\widehat{Vd_{Glu}}} - GFR\cdot\frac{Glucose_{pl}(t)}{\widehat{Vd_{Glu}}} + V_{RGR1}\\
&\quad + V_{RGR2} - Q_{Glu}\cdot\left(\frac{Glucose_{pl}(t)}{\widehat{Vd_{Glu}}} - \frac{Glucose_p(t)}{Vp_{Glu}}\right)\\
\frac{\mathrm{d}Glucose_p}{\mathrm{d}t} &= Q_{Glu}\cdot\left(\frac{Glucose_{pl}(t)}{\widehat{Vd_{Glu}}} - \frac{Glucose_p(t)}{Vp_{Glu}}\right)\\
\frac{\mathrm{d}Glucose_e}{\mathrm{d}t} &= ke_{Glu}\cdot\left(\frac{Glucose_{pl}(t)}{\widehat{Vd_{Glu}}} - Glucose_e(t)\right)\\
\frac{\mathrm{d}Ins_{pl}}{\mathrm{d}t} &= IPRO\cdot\left(\frac{Glucose_e(t)}{GSS}\right)^{IPRG}\cdot(1+S_{incr}\cdot Glucose_{tr}(t)) - CL_{Ins}\cdot\frac{Ins_{pl}(t)}{\widehat{Vd_{Ins}}}\\
\frac{\mathrm{d}Ins_e}{\mathrm{d}t} &= ke_{Ins}\cdot\left(\frac{Ins_{pl}(t)}{\widehat{Vd_{Ins}}} - Ins_e(t)\right)
\end{aligned}
\tag{2.2}
$$

where

$$
\begin{aligned}
M(t) &= \frac{MA}{MD(t)^N + 1}\\
MD(t) &= \frac{time - 12 - 24\cdot floor\left(\frac{time-(MT+12)}{24}\right) - (MT+12)}{MW}\\
\widehat{Vd_{\mathrm{ins}}} &= Vd_{\mathrm{ins}}\cdot\frac{BW}{70}.
\end{aligned}
$$

The final submodel is based on the haemoglobin glycosylation model by Lled-Garca et al. [29]. The module includes a set of transit compartments representing

life cycle of the red blood cells and haemoglobin glycosylation. All parameters are based on the physiological estimates of the respective rates:

$$\frac{\mathrm{d}Hb_1}{\mathrm{d}t} = Kin\cdot\exp(-KG\cdot AG\cdot LSP) - \left(\frac{NC}{LS\cdot AGLS} + KG\cdot AG\right)\cdot Hb_1(t)$$
$$\frac{\mathrm{d}Hb_i}{\mathrm{d}t} = \frac{NC}{LS\cdot AGLS}\cdot Hb_{i-1}(t) - \left(\frac{NC}{LS\cdot AGLS} + KG\cdot AG\right)\cdot Hb_i(t)$$
$$\frac{\mathrm{d}HbA1c_1}{\mathrm{d}t} = Kin\cdot(1-\exp(-KG\cdot AG\cdot LSP)) + KG\cdot AG\cdot Hb_1(t) - \frac{NC}{LS\cdot AGLS}\cdot HbA1c_1(t)$$
$$\frac{\mathrm{d}HbA1c_i}{\mathrm{d}t} = KG\cdot AG\cdot Hb_i(t) - \frac{NC}{LS\cdot AGLS}\cdot(HbA1c_{i-1}(t) - HbA1c_i(t))$$

where $AGLS[-] = (AG/149)^{\gamma}$.

To capture the homeostatic nature of the system, initial conditions were derived from the parameters and baseline states of the variables:

$$Glucose_{pl}(0) = GSS\cdot\widehat{Vd_{Glu}}, \qquad Glucose_e(0) = GSS$$
$$Ins_{pl}(0) = ISS\cdot\widehat{Vd_{Ins}}, \qquad Ins_e(0) = ISS$$
$$IPRO = CL_{Ins}\cdot ISS, \qquad GPRO = GSS\cdot(CL_{Glu} + CL_{Gu}^{Ins}\cdot ISS + GFR) - V_{RGR_{SS}}$$
$$V_{RGR_{SS}} = \left(\frac{Glu_{lum1}(0)}{V_{lumen1}}\right)\cdot\frac{Vmax_{SGLT_2}}{km_{SGLT_2} + Glu_{lum1}(0)/V_{lumen1}} + \left(\frac{Glu_{lum2}(0)}{V_{lumen2}}\right)\cdot\frac{Vmax_{SGLT_1}}{km_{SGLT_1} + Glu_{lum2}(0)/V_{lumen2}}$$
$$Glu_{bl}(0) = \frac{V_{bladder}\cdot Q_{bladder}\cdot Glu_{lum2}(0)}{V_{lumen2}\cdot Q_{urine}}, \qquad Glu_p(0) = GSS\cdot Vp_{Glu}$$
$$Glu_{lum1}(0) = \frac{-b_1 + \sqrt{b_1^2 - 4c_1}}{2}, \qquad Glu_{lum2}(0) = \frac{-b_2 + \sqrt{b_2^2 - 4c_2}}{2}$$
$$Hb_1(0) = Kin\cdot\exp(-KG\cdot AG\cdot LSP)/\left(\frac{NC}{LS\cdot AGLS} + KG\cdot AG\right)$$
$$Hb_i(0) = \frac{NC}{LS\cdot AGLS}\cdot Hb_{i-1}(0)/\left(\frac{NC}{LS\cdot AGLS} + KG\cdot AG\right)$$
$$HbA1c_1(0) = (Hb_1(0)\cdot KG\cdot AG + Kin\cdot(1-\exp(-KG\cdot AG\cdot LSP)))/\left(\frac{NC}{LS\cdot AGLS}\right)$$
$$HbA1c_i(0) = HbA1c_{i-1}(0) + \frac{Hb_i(0)\cdot KG\cdot AG\cdot LS\cdot AGLS}{NC}$$

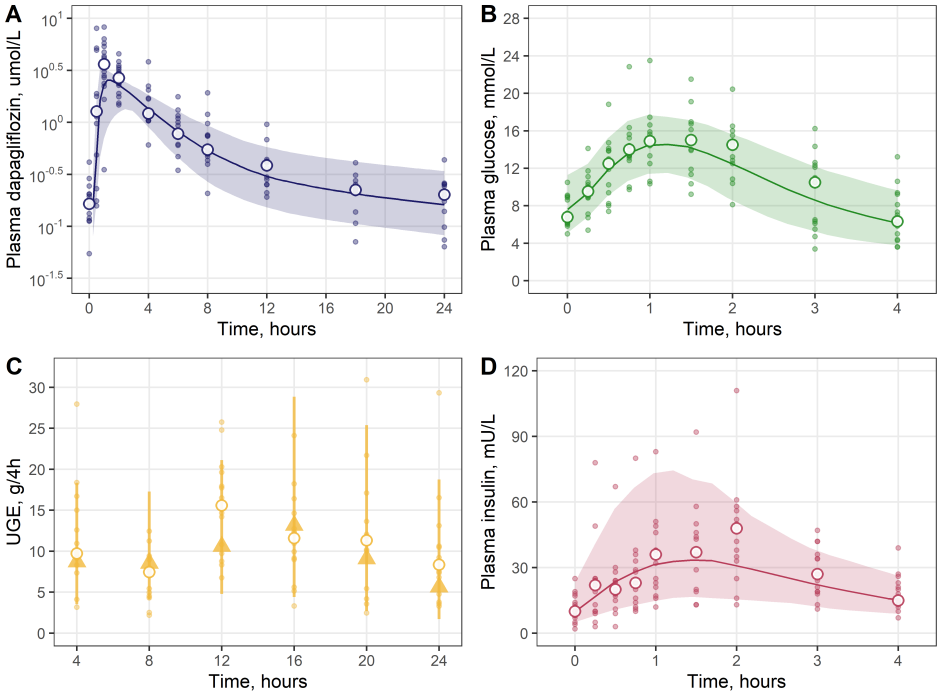

Figure 4. Dapagliflozin PK (A), urinary glucose excretion (C) at day 14, plasma glucose (B) and insulin (D) concentrations at day 13 of the treatment. Dots individual data; empty circles observed median; curves with shaded area (A, B, D) or triangles with error bars (C) predicted median, 5% and 95% percentiles sampled from the between-subject variability.

where

$$b_1 = \left(Km_{SGLT_2} \cdot V_{lumen1} + \frac{V_{lumen1} \cdot Vmax_{SGLT_2}}{Q_{lumen}} - \frac{V_{lumen1} \cdot GFR \cdot GSS}{Q_{lumen}} \right)$$

$$c_1 = -\frac{V_{lumen1}^2 \cdot GFR \cdot GSS \cdot Km_{SGLT_2}}{Q_{lumen}}$$

$$b_2 = \left(Km_{SGLT_1} \cdot V_{lumen2} + \frac{V_{lumen2} \cdot Vmax_{SGLT_1}}{Q_{bladder}} - \frac{V_{lumen2} \cdot Q_{lumen} \cdot Glu_{lum1}(0)}{Q_{bladder} \cdot V_{lumen1}} \right)$$

$$c_2 = -\frac{V_{lumen2}^2 \cdot Q_{lumen} \cdot Glu_{lum1}(0) \cdot Km_{SGLT_1}}{Q_{bladder} \cdot V_{lumen1}}.$$

Initial conditions of the dapagliflozin PK-related variables and food are equal to zero.

As shown in Table 1, more than half of the model parameters were fixed based on the prior knowledge. The remaining parameters were identified based on the subject-level data from the randomized double-blind placebo-controlled Phase 2a dapagliflozin trial [27]. In the study, 47 subjects with T2DM received placebo, 5 mg, 25 mg, or 100 mg of dapagliflozin once per day for 14 days. Time series of daily

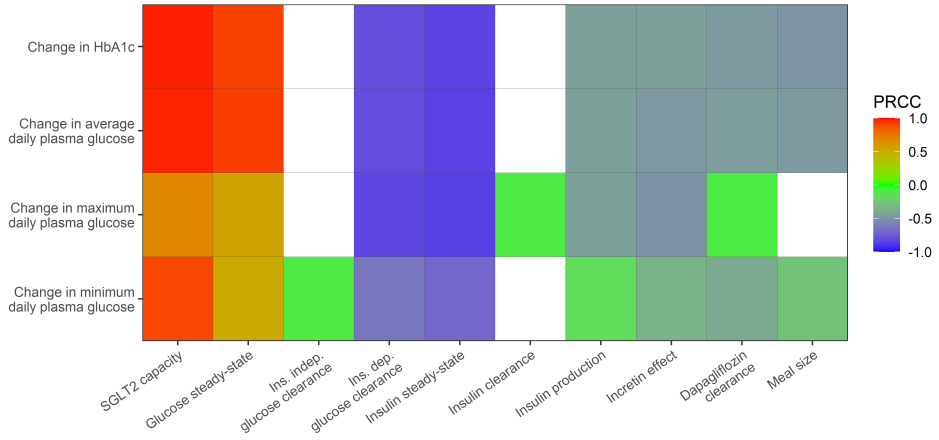

Figure 5. PRCC heatmap. Vertical axis represents the relative change in glycemic control variables at week 30 of daily intake of 10 mg dapagliflozin; PRCC values with p-value $\leqslant 0.05$ are marked with color.

measurements of dapagliflozin, glucose and insulin concentration in plasma as well as glucose amount in urine were available (6631 observations in total). Patients were given meal twice per day and oral glucose tolerance test at day 2 and 13 of the study. Figure 4 shows an exemplary part of the model analysis: comparison of predicted and observed data within a single day of treatment with 100 mg of dapagliflozin.

To highlight the practical application of such platform, a sensitivity analysis was performed in this work by calculating partial rank correlation coefficient (PRCC) for selected variables against the number of parameters [32]. Correlation as a measure of linear relationship between two variables can be defined as follows:

$$\text{corr}_{x_j,y} = \frac{\sum_{i=1}^{N}(x_{ij}-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_{ij}-\bar{x})^2 \cdot \sum_{i=1}^{N}(y_i-\bar{y})^2}}$$

where x_j is the jth parameter, $j = 1,2,,k$, y is the variable, N is the number of samples, $\bar{x}$ and $\bar{y}$ are the sample means. The partial correlation corresponds to the correlation between two residuals $(x_j - \hat{x}_j)$ and $(y - \hat{y})$, where

$$\hat{x}_j = c_0 + \sum_{p=1,p\neq j}^{k} c_p x_p, \qquad \hat{y} = b_0 + \sum_{p=1,p\neq j}^{k} b_p x_p$$

c_0, c_p, b_0, and b_p are the coefficients of multiple linear regression.

Thus, partial rank correlation is the result of partial correlation calculated for the rank-transformed data. Subset of model parameters evaluated via PRCC analysis and respective bounds are summarized in Table 2. Parameter selection was based on their physiological relevance to the gliflozin efficacy and importance in the framework of the glycemic control. Latin hypercube sampling was used to reduce the number of sample size and, thereby, computation load necessary to cover area in the space of parameter values wide enough to obtain consistent results.

Table 2. Parameter bounds for PRCC calculation.

Parameter	Left bound	Right bound	Dimension
Maximum SGLT2 capacity	98.8	125	mmol/h
Glucose steady state	5.38	7.92	mmol/h
Insulin-independent glucose clearance	1.21	2.44	L/h
Insulin-dependent glucose clearance	0.289	0.62	L/(h·mU/L)
Insulin steady state	5.99	14	mU/L
Insulin clearance	67.3	79.6	L/h
Insulin synthesis	1.26	1.6	—
Incretin effect on insulin secretion	0.00763	0.0131	—
Dapagliflozin clearance	12.8	22	L/h
Meal size	0.5	1.5	—

As follows from Fig. 5, parameters responsible for the status of T2DM (glucose and insulin steady states, insulin-dependent glucose clearance) as well as maximum capacity of SGLT2 display the strongest association with dapagliflozin-mediated treatment effects on blood glucose and HbA1c. Dapagliflozin clearance negatively correlates with its efficacy. Increasing the meal size or incretin effect indirectly affect the rate of glucose disappearance through potentiation of insulin secretion, thereby reducing overall treatment benefit.

3. Conclusions

QSP is a modelling technique that combines the elements of systems analysis with physiology and biology. It is a versatile model-informed drug development methodology applied in support of decision-making throughout the process of the development of new pharmaceuticals. The primary goal of a QSP model is to characterize pathophysiological responses to drug interventions quantitatively and mechanistically. However, the spectrum of QSP approaches is vaguely defined and several important limitations exists impairing more extensive application of QSP models. In this study, we developed solutions for the problems of data and knowledge management, choice of the model structure, model credibility assessment, and simulation scenarios, i.e., created a methodology for QSP model development. These methodological concepts are illustrated in a case study of QSP model development in T2DM applied to identify patient subpopulation most responsive to the gliflozin treatment.

Acknowledgment: The author thanks Gennady Bocharov and Kirill Peskov for mentorship, guidance and feedback throughout the scientific career, and Cristina Leon for invaluable help with the arrangement of this article.

References

1. American Diabetes Association, 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes2020. *Dia. Care* **43** (2020), S14–S31.

2. J. P. F. Bai, B. J. Schmidt, K. G. Gadkar, V. Damian, J. C. Earp, C. Friedrich, P. H. van der Graaf, R. Madabushi, C. J. Musante, K. Naik, M. Rogge, and H. Zhu, FDA-Industry Scientific Exchange

on assessing quantitative systems pharmacology models in clinical drug development: a meeting report, summary of challenges/gaps, and future perspective. *AAPS J.* **23** (2021), 60.

3. Y. Bard, *Nonlinear Parameter Estimation*. Academic Press, New York, 1974.

4. K. T. Baron, S. Macha, U. C. Broedl, V. Nock, S. Retlich, and M. Riggs, Population pharmacokinetics and exposure-response (efficacy and safety/tolerability) of empagliflozin in patients with type 2 diabetes. *Diabetes Ther.* **7** (2016), 455–471.

5. R. J. Bauer, NONMEM tutorial Part II: Estimation methods and advanced examples. *CPT Pharmacometrics Syst. Pharmacol.* **8** (2019), 538–556.

6. K. W. Beach, A theoretical model to predict the behavior of glycosylated hemoglobin levels. *J. Theor. Biol.* **81** (1979), 547–561.

7. L. Z. Benet and M. Rowland, Pharmacometrics: A new journal section. *Journal of Pharmacokinetics and Biopharmaceutics* **10** (1982), 349350.

8. R. N. Bergman, Y. Z. Ider, C. R. Bowden, and C. Cobelli, Quantitative estimation of insulin sensitivity. *Am. J. Physiol.* **236** (1979), E667–677.

9. E. L. Bradshaw, M. E. Spilker, R. Zang, L. Bansal, H. He, R. D. O. Jones, K. Le, M. Penney, E. Schuck, B. Topp, A. Tsai, C. Xu, M. J. M. A. Nijsen, and J. R. Chan, Applications of quantitative systems pharmacology in model-informed drug discovery: perspective on impact and opportunities. *CPT Pharmacometrics Syst Pharmacol.* **8** (2019), 777–791.

10. A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*. Wiley-Blackwell, Weinheim, Germany, 2012.

11. A. De Gaetano and O. Arino, Mathematical modelling of the intravenous glucose tolerance test. *J. Math. Biol.* **40** (2000), 136–168.

12. R. A. DeFronzo, E. Ferrannini, L. Groop, R. R. Henry, W. H. Herman, J. J. Holst, F. B. Hu, C. R. Kahn, I. Raz, G. I. Shulman, D. C. Simonson, M. A. Testa, and R. Weiss, Type 2 diabetes mellitus. *Nat. Rev. Dis. Primers.* **1** (2015), 15019.

13. D. Devineni, D. Polidori, C. R. Curtin, J. Murphy, S.-S. Wang, H. Stieltjes, and E. Wajs, Pharmacokinetics and pharmacodynamics of once- and twice-daily multiple-doses of canagliflozin, a selective inhibitor of sodium glucose co-transporter 2, in healthy participants. *Int. J. Clin. Pharmacol. Ther.* **53** (2015), 438–446.

14. B. Efron and R. Tibshirani, Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statist. Sci.* **1** (1986), 54–75.

15. S. Ermakov, B. J. Schmidt, C. J. Musante, and C. J. Thalhauser, A survey of software tool utilization and capabilities for quantitative systems pharmacology: what we have and what we need. *CPT Pharmacometrics Syst. Pharmacol.* **8** (2019), 62–76.

16. A. Field, J. Miles, and Z. Field, *Discovering statistics using R*. Sage, Los Angeles–London–New Delhi–Singapore–Washington, DC, 2012.

17. C. Ghezzi, D. D. F. Loo, and E. M. Wright, Physiology of renal glucose handling via SGLT1, SGLT2 and GLUT2. *Diabetologia* **61** (2018), 2087–2097.

18. K. M. Hallow, P. J. Greasley, G. Helmlinger, L. Chu, H. J. Heerspink, and D. W. Boulton, Evaluation of renal and cardiovascular protection mechanisms of SGLT2 inhibitors: model-based analysis of clinical data. *American Journal of Physiology-Renal Physiology* **315** (2018), F1295–F1306.

19. B. Hamrn, E. Bjrk, M. Sunzel, and M. Karlsson, Models for plasma glucose, HbA1c, and hemoglobin interrelationships in patients with type 2 diabetes following tesaglitazar treatment. *Clin. Pharmacol. Ther.* **84** (2008), 228–235.

20. G. Helmlinger, N. Al-Huniti, S. Aksenov, K. Peskov, K. M. Hallow, L. Chu, D. Boulton, U. Eriksson, B. Hamrn, C. Lambert, E. Masson, H. Tomkinson, and D. Stanski, Drug-disease modeling in the pharmaceutical industry—where mechanistic systems pharmacology and statistical pharmacometrics meet.*European Journal of Pharmaceutical Sciences.* **109** (2017), S39–S46.

21. P. J. Higgins and H. F. Bunn, Kinetic analysis of the nonenzymatic glycosylation of hemoglobin. *J. Biol. Chem.* **256** (1981), 5204–5208.

22. P. M. Jauslin, N. Frey, and M. O. Karlsson, Modeling of 24-hour glucose and insulin profiles of patients with type 2 diabetes. *J. Clin. Pharmacol.* **51** (2011), 153–164.

23. P. M. Jauslin, M. O. Karlsson, and N. Frey, Identification of the mechanism of action of a glucokinase activator from oral glucose tolerance test data in type 2 diabetic patients based on an integrated glucose-insulin model. *J. Clin. Pharmacol.* **52** (2012), 1861–1871.

24. P. M. Jauslin, H. E. Silber, N. Frey, R. Gieschke, U. S. H. Simonsson, K. Jorga, and M. O. Karlsson, An integrated glucose-insulin model to describe oral glucose tolerance test data in type 2 diabetics. *J. Clin. Pharmacol.* **47** (2007), 1244–1255.

25. C. K. Johnston, R. J. Eudy-Byrne, A. Elmokadem, V. Nock, J. Marquard, N. Soleymanlou, M. M. Riggs, K.-H. Liesenfeld, A model-informed drug development (MIDD) approach for a low dose of empagliflozin in patients with type 1 diabetes. *Pharmaceutics.* **13** (2021), 485.

26. V. A. Karpelev, Yu. I. Filippov, Yu. V. Tarasov, M. D. Boyarsky, A. Yu. Mayorov, M. V. Shestakova, and I. I. Dedov, Mathematical modeling of the blood glucose regulation system in diabetes mellitus patients. *Annals RAMS.* **70** (2015), 549–560.

27. B. Komoroski, N. Vachharajani, Y. Feng, L. Li, D. Kornhauser, and M. Pfister, Dapagliflozin, a novel, selective SGLT2 inhibitor, improved glycemic control over 2 weeks in patients with type 2 diabetes mellitus. *Clin. Pharmacol. Ther.* **85** (2009), 513–519.

28. M. Lavielle and F. Mentr, Estimation of population pharmacokinetic parameters of saquinavir in HIV patients with the MONOLIX software. *J. Pharmacokinet. Pharmacodyn.* **34** (2007), 229–249.

29. R. Lled-Garca, N. A. Mazer, and M. O. Karlsson, A semi-mechanistic model of the relationship between average glucose and HbA1c in healthy and diabetic subjects. *J. Pharmacokinet. Pharmacodyn.* **40** (2013), 129–142.

30. R. Madabushi, P. Seo, L. Zhao, M. Tegenge, and H. Zhu, Review: role of model-informed drug development approaches in the lifecycle of drug development and regulatory decision-making. *Pharm Res.* **39** (2022), 1669–1680.

31. R. Malka, D. M. Nathan, and J. M. Higgins, Mechanistic modeling of hemoglobin glycation and red blood cell kinetics enables personalized diabetes monitoring. *Sci. Transl. Med.* **8** (2016), 359ra130.

32. S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner, A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology* **254** (2008), 178–196.

33. J. Melin, J. Parkinson, and D. Rekic, Dapagliflozin pharmacokinetics is similar in adults with type 1 and type 2 diabetes mellitus, Abstr 8427. *pp. 27*, Montreux, Switzerland, 2018.

34. P. A. Milligan, M. J. Brown, B. Marchant, S. W. Martin, P. H. van der Graaf, N. Benson, G. Nucci, D. J. Nichols, R. A. Boyd, J. W. Mandema, S. Krishnaswami, S. Zwillich, D. Gruben, R. J. Anziano, T. C. Stock, and R. L. Lalonde, Model-based drug development: a rational approach to efficiently accelerate drug development. *Clin. Pharmacol. Ther.* **93** (2013), 502–514.

35. J. B. Mller, R. V. Overgaard, M. C. Kjellsson, N. R. Kristensen, S. Klim, S. H. Ingwersen, and M. O. Karlsson, Longitudinal modeling of the eelationship between mean plasma glucose and HbA1c following antidiabetic treatments. *CPT Pharmacometrics Syst. Pharmacol.* **2** (2013), e82.

36. H. B. Mortensen, A. Vlund, and C. Christophersen, Glucosylation of human haemoglobin A. Dynamic variation in HbA1c described by a biokinetic model. *Clin. Chim. Acta.* **136** (1984), 75–81.

37. T. H. T. Nguyen, M. Mouksassi, N. Holford, N. Al-Huniti, I. Freedman, A. C. Hooker, J. John, M. O. Karlsson, D. R. Mould, J. J. Prez Ruixo, E. L. Plan, R. Savic, J. G. C. van Hasselt, B. Weber, C. Zhou, E. Comets, and F. Mentr, For the Model Evaluation Group of the International Society of Pharmacometrics (ISoP) Best Practice Committee, Model evaluation of continuous data pharmacometric models: metrics and graphics. *CPT Pharmacometrics Syst. Pharmacol.* **6** (2017), 87–109.

38. M. J. M. A. Nijsen, F. Wu, L. Bansal, E. Bradshaw-Pierce, J. R. Chan, B. M. Liederer, J. T. Mettetal, P. Schroeder, E. Schuck, A. Tsai, C. Xu, A. Chimalakonda, K. Le, M. Penney, B. Topp, A. Yamada, and M. E. Spilker, Preclinical QSP modeling in the pharmaceutical industry: an IQ consortium survey examining the current landscape: preclinical QSP in pharmaceutical industry. *CPT Pharmacometrics Syst. Pharmacol.* **7** (2018), 135–146.

39. S. M. Osterman-Golkar and H. W. Vesper, Assessment of the relationship between glucose and A1c using kinetic modeling. *J. Diabetes Complications.* **20** (2006), 285–294.

40. M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hrbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher, The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372** (2021), n71.

41. J. Parkinson, B. Hamrn, M. C. Kjellsson, and S. Skrtic, Application of the integrated glucose-insulin model for cross-study characterization of T2DM patients on metformin background treatment. *Brit. J. Clin. Pharmacol.* **82** (2016), 1613–1624.

42. C. C. Peck, W. H. Barr, L. Z. Benet, J. Collins, R. E. Desjardins, D. E. Furst, J. G. Harter, G. Levy, T. Ludden, and J. H. Rodman, Opportunities for integration of pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development. *J. Clin. Pharmacol.* **34** (1994), 111–119.

43. B. A. Perkins, N. Soleymanlou, J. Rosenstock, J. S. Skyler, L. M. Laffel, K.-H. Liesenfeld, D. Neubacher, M. M. Riggs, C. K. Johnston, R. J. Eudy-Byrne, A. Elmokadem, J. T. George, J. Marquard, and V. Nock, Low-dose empagliflozin as adjunct-to-insulin therapy in type 1 diabetes: A valid modelling and simulation analysis to confirm efficacy. *Diabetes Obes. Metab.* **22** (2020), 427–433.

44. M. M. Riggs, L. J. Seman, A. Staab, T. R. MacGregor, W. Gillespie, M. R. Gastonguay, H. J. Woerle, and S. Macha, Exposure-response modelling for empagliflozin, a sodium glucose co-transporter 2 (SGLT2) inhibitor, in patients with type 2 diabetes. *Brit. J. Clin. Pharmacol.* **78** (2014), 1407–1418.

45. M. M. Riggs, A. Staab, L. Seman, T. R. MacGregor, T. T. Bergsma, M. R. Gastonguay, and S. Macha, Population pharmacokinetics of empagliflozin, a sodium glucose cotransporter 2 inhibitor, in patients with type 2 diabetes. *J. Clin. Pharmacol.* **53** (2013), 1028–1038.

46. P. V. Rder, B. Wu, Y. Liu, and W. Han, Pancreatic regulation of glucose homeostasis. *Exp. Mol. Med.* **48** (2016), e219–e219.

47. R. M. Rge, S. Klim, S. H. Ingwersen, M. C. Kjellsson, and N. R. Kristensen, The effects of a GLP-1 analog on glucose homeostasis in type 2 diabetes mellitus quantified by an integrated glucose insulin model. *CPT Pharmacometrics Syst. Pharmacol.* **4** (2015), e00011.

48. M. Shah, L. Stolbov, T. Yakovleva, W. Tang, V. Sokolov, R. C. Penland, D. Boulton, and J. Parkinson, A model-based approach to investigating the relationship between glucoseinsulin dynamics and dapagliflozin treatment effect in patients with type 2 diabetes. *Diabetes Obes. Metab.* **23** (2021), 991–1000.

49. L. B. Sheiner, Computer-aided long-term anticoagulation therapy. *Comput. Biomed. Res.* **2** (1969), 507–518.

50. L. B. Sheiner, Learning versus confirming in clinical drug development, *Clin. Pharmacol. Ther.* **61** (1997), 275–291.

51. L. B. Sheiner and S. L. Beal, Evaluation of methods for estimating population pharmacokinetics parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. *J. Pharmacokinet. Biopharm.* **8** (1980), 553–571.

52. H. E. Silber, N. Frey, and M. O. Karlsson, An integrated glucose-insulin model to describe oral glucose tolerance test data in healthy volunteers. *J. Clin. Pharmacol.* **50** (2010), 246–256.

53. H. E. Silber, P. M. Jauslin, N. Frey, R. Gieschke, U. S. H. Simonsson, and M. O. Karlsson, An integrated model for glucose and insulin regulation in healthy volunteers and type 2 diabetic patients following intravenous glucose provocations. *J. Clin. Pharmacol.* **47** (2007), 1159–1171.

54. D. J. Venzon and S. H. Moolgavkar, A Method for computing profile-likelihood-based confidence intervals. *Applied Statistics* **37** (1988), 87.

55. W. de Winter, J. DeJongh, T. Post, B. Ploeger, R. Urquhart, I. Moules, D. Eckland, and M. Danhof, A mechanism-based disease progression model for comparison of long-term effects of pioglitazone, metformin and gliclazide on disease processes underlying type 2 diabetes mellitus. *J. Pharmacokinet. Pharmacodyn.* **33** (2006), 313–343.

56. W. de Winter, A. Dunne, X.W. de Trixhe, D. Devineni, C.-H. Hsu, J. Pinheiro, and D. Polidori, Dynamic population pharmacokinetic-pharmacodynamic modelling and simulation supports similar efficacy in glycosylated haemoglobin response with once or twice-daily dosing of canagliflozin. *Brit. J. Clin. Pharmacol.* **83** (2017), 1072–1081.

57. T. Yakovleva, V. Sokolov, L. Chu, W. Tang, P. J. Greasley, H. Peilot Sjgren, S. Johansson, K. Peskov, G. Helmlinger, D. W. Boulton, and R. C. Penland, Comparison of the urinary glucose excretion contributions of SGLT2 and SGLT1: A quantitative systems pharmacology analysis in healthy individuals and patients with type 2 diabetes treated with SGLT2 inhibitors. *Diabetes Obes. Metab.* **21** (2019), 2684–2693.

58. NIH Workgroup, *Quantitative and Systems Pharmacology in the Post-genomic Era: New Approaches to Discovering Drugs and Understanding Therapeutic Mechanisms*, 2011. (Paper presented at the QSP Workshop Group).