Grey-box Modelling of Pharmacokinetic /Pharmacodynamic Systems

Christoffer W. Tornøe[1],*, Judith L. Jacobsen[2], Oluf Pedersen[3], Torben Hansen[3,4], and Henrik Madsen[1]

Grey-box pharmacokinetic/pharmacodynamic (PK/PD) modelling is presented as a promising way of modelling PK/PD systems. The concept behind grey-box modelling is based on combining physiological knowledge along with information from data in the estimation of model parameters. Grey-box modelling consists of using stochastic differential equations (SDEs) where the stochastic term in the differential equations represents unknown or incorrectly modelled dynamics of the system. The methodology behind the grey-box PK/PD modelling framework for systematic model improvement is illustrated using simulated data and furthermore applied to Bergman's minimal model of glucose kinetics using clinical data from an intravenous glucose tolerance test (IVGTT). The grey-box estimates of the stochastic system noise parameters indicate that the glucose minimal model is too simple and should preferably be revised in order to describe the complicated in vivo system of insulin and glucose following an IVGTT.

KEY WORDS: grey-box PK/PD modelling; stochastic differential equations; systematic model improvement; glucose minimal model; IVGTT.

INTRODUCTION

Grey-box pharmacokinetic/pharmacodynamic (PK/PD) modelling combines a known or proposed model structure with a stochastic term in the differential equations representing disturbances, inputs to the system which are not measured, and unmodelled dynamics of the system. This

---

[1]Informatics and Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark.
[2]Department of Biostatistics, Novo Nordisk A/S, Bagsværd, Denmark.
[3]Steno Diabetes Center, Copenhagen, Denmark.
[4]Department of Clinical Biochemistry, Rigshospitalet, Copenhagen, Denmark.
*To whom correspondence should be addressed: Ferring Pharmaceuticals A/S, Kay Fiskers Plads 11, DK-2300, Copenhagen S, Denmark. Tel:+45-2878-7453; fax:+45-2817-6453; e-mail: christoffer.tornoe@ferring.com

setup facilitates a more rigorous statistical approach for parameter estimation where the residual error is decomposed into measurement and system error (1). The advantage of using grey-box PK/PD models is that physiological knowledge is combined with information from data. This feature makes it a very attractive tool for modelling PK/PD systems where the system is not fully understood or cannot be explicitly modelled. The setup can furthermore be used as a suitable framework for systematic model improvement (2) illustrated by a simulated example of an indirect response model (3) and by Bergman's minimal model of glucose kinetics (GMM) (4, 5) using clinical data from an intravenous glucose tolerance test (IVGTT).

THEORY

Stochastic Differential Equations

Grey-box models are stochastic state space models consisting of stochastic differential equations (SDEs) describing the dynamics of the states of the system and discrete time measurement equations describing the observations as a function of the states. Grey-box models thereby accommodate random effects by decomposing the noise into a stochastic system or process noise term representing unknown or incorrectly modelled dynamics of the system and a stochastic measurement noise term accounting for uncorrelated errors such as assay error (1).

The relationship between input and output signals in a dynamical system is conveniently described by a state space model. A state space model is an internal parametric representation between input and output which in a continuous time formulation enables a direct physiological meaning of the parameters. Since the structural information of the system is formulated in continuous time and the data is observed at discrete time points, the continuous-discrete time stochastic state space model consisting of a continuous time system equation and a discrete time observation equation is defined as

$$dx = g(\theta, x, d)dt + \sigma_w(\theta, d)dw_t, \qquad (1)$$

$$y_j = f(\theta, x, d) + e_j, \qquad (2)$$

where x is the state vector, the vector d contains the inputs to the system, θ is the parameter vector, and t is the time variable. The deterministic

function $\boldsymbol{g}(\cdot)$ is called the drift term, $\boldsymbol{\sigma}_w(\cdot)$ is the diffusion term while $\boldsymbol{w}_t$ is a standard Wiener process also referred to as Brownian motion (6). The standard Wiener process $\boldsymbol{w}_t$ is a non-stationary stochastic process with mutually independent (orthogonal) increments which are normally distributed with mean and variance

$$E[\boldsymbol{w}_{t_2} - \boldsymbol{w}_{t_1}] = \boldsymbol{0}, \tag{3}$$

$$V[\boldsymbol{w}_{t_2} - \boldsymbol{w}_{t_1}] = |t_2 - t_1|\boldsymbol{I}, \tag{4}$$

where $\boldsymbol{I}$ is the identity matrix.

By setting the diffusion term $\boldsymbol{\sigma}_w(\cdot)$ in (1) equal to zero, an ordinary differential equation (ODE) is obtained. The transformation from ordinary to stochastic differential equations does therefore not affect the drift term $\boldsymbol{g}(\cdot)$ and the physiological interpretation of the parameters is preserved.

The vector of measurements at time t_j, denoted by $\boldsymbol{y}_j$, is modelled as a function of states and inputs as described by $\boldsymbol{f}(\cdot)$ while the measurement error terms $\boldsymbol{e}_j$ are independent and identically distributed (iid) with mean zero and variance $\Sigma = \boldsymbol{\sigma\sigma}^T$. Finally, we assume that $\boldsymbol{w}_t$ and $\boldsymbol{e}_j$ are independent.

Maximum Likelihood Estimation in SDEs

The maximum likelihood (ML) method used for estimating the parameters in grey-box models is described in the following (see (7–9) for further details). Maximum likelihood estimation is based on maximizing the likelihood function in term of the joint distribution of the observations $\mathscr{Y}_N = [\boldsymbol{y}_N, \boldsymbol{y}_{N-1}, \ldots, \boldsymbol{y}_0]$ given the parameter vector $\boldsymbol{\theta}$ (9).

$$L(\boldsymbol{\theta}, \mathscr{Y}_N) = p(\mathscr{Y}_N|\boldsymbol{\theta}) = \left(\prod_{j=1}^{N} p(\boldsymbol{y}_j|\mathscr{Y}_{j-1}, \boldsymbol{\theta})\right) p(\boldsymbol{y}_0|\boldsymbol{\theta}) \tag{5}$$

where the Markov property $P(A \cap B) = P(A|B)P(B)$ has been applied to express the likelihood function as a product of marginal conditional probability densities. Under the assumption that the conditional density function is Gaussian (which is true for linear models), the following equations entirely characterize the one-step conditional density $p(\boldsymbol{y}_j|\mathscr{Y}_{j-1}, \boldsymbol{\theta})$ in (5), i.e.,

$$\hat{\boldsymbol{y}}_{j|j-1} = E[\boldsymbol{y}_j | \mathscr{Y}_{j-1}, \boldsymbol{\theta}] \tag{6}$$

$$\boldsymbol{R}_{j|j-1} = V[\boldsymbol{y}_j | \mathscr{Y}_{j-1}, \boldsymbol{\theta}] \tag{7}$$

By introducing the one-step prediction error ϵ_j

$$\epsilon_j = \boldsymbol{y}_j - \hat{\boldsymbol{y}}_{j|j-1} \tag{8}$$

the likelihood function can then be written as

$$L(\boldsymbol{\theta}; \mathscr{Y}_N) = \left(\prod_{j=1}^{N} \frac{\exp\left(-\frac{1}{2}\epsilon_j^T \boldsymbol{R}_{j|j-1}^{-1} \epsilon_j\right)}{\sqrt{\det(\boldsymbol{R}_{j|j-1})}(\sqrt{2\pi})^l} \right) p(\boldsymbol{y}_0 | \boldsymbol{\theta}), \tag{9}$$

where l is the dimension of $\boldsymbol{y}_j$.

If the likelihood function is further conditioned on $\boldsymbol{y}_0$ and by taking the logarithm of (9), the following log-likelihood function is obtained.

$$\ln\left(L(\boldsymbol{\theta}; \mathscr{Y}_N | \boldsymbol{y}_0)\right) = -\frac{1}{2} Nl \ln(2\pi) - \frac{1}{2} \sum_{j=1}^{N} \left(\ln(\det(\boldsymbol{R}_{j|j-1})) + \epsilon_j^T \boldsymbol{R}_{j|j-1}^{-1} \epsilon_j \right) \tag{10}$$

The one-step prediction error ϵ_j and the conditional covariance $\boldsymbol{R}_{j|j-1}$ can be calculated for given parameters $\boldsymbol{\theta}$ and initial conditions $\boldsymbol{x}_0$ by the use of the Kalman filter (KF) or the extended Kalman filter (EKF) for linear and non-linear models, respectively. The ML estimate of the parameters $\boldsymbol{\theta}$ is now obtained as

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \{-\ln(L(\boldsymbol{\theta}; \mathscr{Y}_N | \boldsymbol{y}_0))\} \tag{11}$$

The maximum likelihood estimates are asymptotically normally distributed with mean $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{D}$ found from the Fisher information matrix (9). An approximation of $\boldsymbol{D}$ can be found by

$$\boldsymbol{D}(\hat{\boldsymbol{\theta}}) \simeq \boldsymbol{H}^{-1}, \tag{12}$$

where H is the Hessian calculated as the second derivative of the log-likelihood function in (10) at the obtained parameter estimates.

State Filtering

The ML estimation of parameters in SDEs involves the use of state filtering. The EKF is a set of mathematical equations that provides an efficient recursive approach to approximate the conditional densities with Gaussian distributions (6). The conditional densities are thus completely characterized by the one-step prediction error ϵ_j and the associated one-step prediction covariance matrix $R_{j|j-1}$.

Given the parameters θ, initial states $\hat{x}_{t|t_0} = x_0$, and initial state covariance $\hat{P}_{t|t_0}$, the *prediction equations* of the EKF are the optimal (minimum variance) prediction of the state mean and state covariance, i.e.,

$$\frac{d\hat{x}_{t|j-1}}{dt} = g(\theta, \hat{x}_{t|j-1}, d), \tag{13}$$

$$\frac{dP_{t|j-1}}{dt} = AP_{t|j-1} + P_{t|j-1}A^T + \sigma_w \sigma_w^T, \tag{14}$$

where $\hat{x}_{t|j-1}$ is the one-step state prediction and $P_{t|j-1}$ is the associated state covariance. If the drift term $g(\cdot)$ is a non-linear function in x, a local first-order Taylor expansion of $g(\cdot)$ at each sampling time is used to describe the dynamics of the system. The matrix A is thus approximated by

$$A = \left.\frac{\partial g}{\partial x}\right|_{x=\hat{x}_{t|j-1}}. \tag{15}$$

The *output prediction equations* of the observation $\hat{y}_{j|j-1}$ and the associated covariance $R_{j|j-1}$ are calculated by

$$\hat{y}_{j|j-1} = C\hat{x}_{j|j-1}, \tag{16}$$

$$R_{j|j-1} = CP_{j|j-1}C^T + \sigma\sigma^T, \tag{17}$$

where $\boldsymbol{\sigma}$ is the standard deviation of the measurement error and $\boldsymbol{C}$ is the linearization of the measurement equation, i.e.,

$$\boldsymbol{C} = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\hat{\boldsymbol{x}}_{j|j-1}}. \tag{18}$$

Finally, the *update equations* of the EKF are

$$\hat{\boldsymbol{x}}_{j|j} = \hat{\boldsymbol{x}}_{j|j-1} + \boldsymbol{K}_j(\boldsymbol{y}_j - \hat{\boldsymbol{y}}_{j|j-1}), \tag{19}$$

$$\boldsymbol{P}_{j|j} = \boldsymbol{P}_{j|j-1} - \boldsymbol{K}_j \boldsymbol{R}_{j|j-1} \boldsymbol{K}_j^T, \tag{20}$$

$$\boldsymbol{K}_j = \boldsymbol{P}_{j|j-1} \boldsymbol{C}^T \boldsymbol{R}_{j|j-1}^{-1}, \tag{21}$$

where $\hat{\boldsymbol{x}}_{j|j}$ is the updated state prediction, $\boldsymbol{P}_{j|j}$ is the associated covariance, and $\boldsymbol{K}_j$ is the Kalman gain.

A more detailed description of the theory behind stochastic differential equations can be found in (6, 10) while their application to pharmacokinetic/ pharmacodynamic (PK/PD) modelling is further discussed in (7, 11, 12, 13).

EXPERIMENT

Intravenous Glucose Tolerance Test

The IVGTT experiment consists of injecting a bolus of glucose into the bloodstream thus inducing an increase of the plasma insulin concentration secreted by the pancreas. The IVGTT was performed on the offspring and spouses of 70 verified type II diabetes mellitus subjects. Prior to the participation in the study informed consent was obtained from all subjects. The study was performed at Steno Diabetic Center and approved by the Ethical Committee of Copenhagen and was in accordance with the principles of the Declaration of Helsinki.

About 219 normal glucose tolerant (NGT) and 25 impaired glucose tolerant (IGT) subjects underwent a tolbutamide modified, frequently sampled IVGTT. After 12 hr of fasting, venous blood samples were drawn in triplicate at −10, −5, and 0 min before the IVGTT and at 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 16, 19, 22, 23, 24, 25, 27, 30, 35, 40, 50, 60, 70, 80, 90,

100, 120, 140, 160, and 180 min after the IV bolus for analysis of plasma insulin, glucose, and C-peptide concentrations. At $t = 0$ min, glucose was injected intravenously in the contralateral antecubital vein over a period of 1 min (0.3 g/(kg BW) of 50% glucose). At $t = 20$ min, a bolus of 3 mg tolbutamide/(kg BW) was injected over 5 s to elicit a secondary pancreatic beta-cell response, i.e., an insulin burst. This ensures that as much of the dynamics of the insulin/glucose system as possible is present in the measured data. For IGT subjects, tolbutamide is specifically injected to raise the plasma insulin level above the baseline value.

RESULTS AND DISCUSSION

Simulated Example

Let us first use a simulated example to illustrate the methodology behind the grey-box PK/PD modelling framework for uncovering model deficiencies through iterative model development.

The PK data is generated by a one-compartment model with an IV constant infusion of 400 mg given over 1 hr. The PD data is simulated from an indirect response model (3) where the plasma drug concentration stimulated the loss of response, i.e.,

$$\frac{dA}{dt} = -\frac{CL}{V} A, \tag{22}$$

$$\frac{dR}{dt} = K_{in} - k_{out}\left(1 + \frac{E_{max}(A/V)^\gamma}{EC_{50}^\gamma + (A/V)^\gamma}\right)R, \tag{23}$$

where A is the state variable for the amount of drug in the plasma and R is the state variable for the PD response. The data is simulated using the following parameter values: clearance CL $= 1$ l/h, volume of distribution $V = 25$ l, rate-constant for production of response $K_{in} = 1$ mg/l/hr, first-order rate-constant for loss of response $k_{out} = 0.1$ hr^{-1} maximal stimulating effect $E_{max} = 2$, drug concentration causing 50% of maximal stimulation $EC_{50} = 5$ mg/l, and sigmoidicity factor $\gamma = 5$.

By assuming stationarity before drug administration, the initial conditions for the system of differential equations in (22)–(23) can be specified by

$$A_0 = 0, \tag{24}$$

$$R_0 = \frac{K_{in}}{k_{out}}, \tag{25}$$

where A_0 is the amount of drug in the plasma at $t = 0$ hr while R_0 is the baseline response concentration at $t = 0$ hr.

The measurement error models used to simulate the PK/PD data are

$$\log PK_j = \log \frac{A_j}{V} + e_j^{\text{PK}}, \tag{26}$$

$$PD_j = R_j + e_j^{PD}, \tag{27}$$

where PK and PD are the observed plasma drug and response concentrations, respectively, observed at discrete time instants t_j. The measurement error terms e_j^{PK} and e_j^{PD} are obtained by sampling from two normal densities with zero mean and variance $\sigma_{\text{PK}}^2 = 0.15^2$ and $\sigma_{\text{PD}}^2 = 0.3^2$. The methodology described in (2) for using a grey-box modelling cycle for systematic model improvement is used in the following. The first step in the modelling cycle is to use prior physiological knowledge to derive an initial model structure in the form of an ordinary differential equation (ODE) model and translate it into a stochastic state-space model by adding a diffusion term. Assuming that the initial model structure corresponding to (22)–(23) is correct except for the functional relationship between the PK and PD data, the first grey-box PK/PD modelling cycle iteration model is

$$\begin{bmatrix} dA \\ dR \end{bmatrix} = \begin{bmatrix} -\frac{CL}{V} A \\ K_{in} - k_{\text{out}} \left(1 + \frac{A/V}{EC_{50}+A/V} \right) R \end{bmatrix} dt + \begin{bmatrix} \sigma_A & 0 \\ 0 & \sigma_R \end{bmatrix} d\mathbf{w}_t, \tag{28}$$

where the functional relation between the plasma drug concentration and the PD response is an $E_{\max}$ model since the true structure is assumed unknown. A diagonal parameterization of the diffusion term in the system equation is used to allow model deficiencies to be located if the model is insufficient. By fully conditioning on the initial response observation, the baseline response concentration R_0 is fixed at the observed concentration at $t = 0$ while disregarding the observation at $t = 0$. The rate-constant for production of response K_{in} can thereby be substituted with $k_{\text{out}}R_0$. The measurement equations used throughout the iterative model development are the same as those used for simulating the data, i.e., (26)–(27).

The unknown parameters in the first model iteration described by (28) are estimated using maximum likelihood (ML) estimation

Table I. Parameter estimates for the three grey-box PK/PD modelling cycle iterations

Iteration			1		2		3	
Parameter	Unit	True	Estimate	RSE (%)	Estimate	RSE (%)	Estimate	RSE (%)
E_0	[mg/L]	–	–	–	2.68	40.4	–	–
CL	[L/hr]	1.0	1.04	3.58	1.04	3.52	1.04	3.71
V	[L]	25.0	25.8	4.36	25.9	6.18	25.9	4.96
k_{out}	[1/hr]	0.1	0.0736	51.3	0.0527	41.9	0.0846	9.80
E_{max}	[–]	2.0	–	–	–	–	2.09	6.50
EC_{50}	[mg/L]	5.0	3.49	59.8	–	–	4.82	7.21
γ	[–]	5.0	–	–	–	–	6.92	27.6
σ_A	[mg/$\sqrt{hr}$]	0.0	0.00	193	0.00	139	0.00	766
σ_R	[mg/L $\sqrt{hr}$]	0.0	0.759	22.0	0.00	>1000	0.00	>1000
σ_E	[1/$\sqrt{hr}$]	0.0	–	–	0.239	42.6	–	–
σ_{PK}	[%]	15.0	13.5	18.6	13.5	18.2	13.5	18.1
σ_{PD}	[mg/L]	0.30	0.00	>1000	0.326	25.6	0.244	18.7

(see Table I). On a 5% significance level, the only diffusion parameter which is insignificant is σ_A. Since σ_R is significant, it indicates that the PD system equation in (28) is incorrectly specified since no diffusion term was included in the simulated data. In practice, a significant or large value of the diffusion term points to a state variable where there is potential room for model improvement.

The predicted concentration–time profiles are shown in Fig. 1 along with the observed concentrations. The PK model fits the observed data nicely while the suggested PD model clearly does not capture the true dynamics of the PD system when comparing the observed and predicted concentrations.
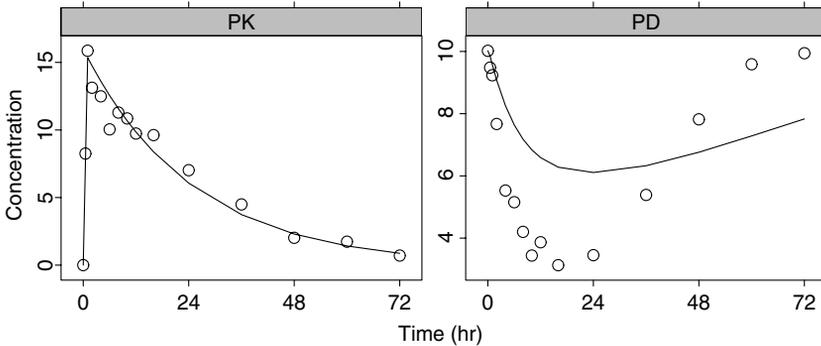

Fig. 1. Plot of observed (circles) and predicted (solid lines) concentrations from the first grey-box model iteration. PK (Left) and PD (Right).

Instead of trying every possible combination for the functional relationship between the PK and PD data, the second step in the grey-box PK/PD modelling cycle iteration consists of re-formulating the model in (28) with an additional state variable E representing the effect model. The drift term in the state equation for E is set equal to zero in dE/dt thereby assuming that E is a constant while any variations is adopted by the diffusion term parameterized by σ_E. The system equations for the second model iteration can thereby be written as

$$
\begin{bmatrix} dA \\ dR \\ dE \end{bmatrix} = \begin{bmatrix} -\frac{CL}{V}A \\ K_{\text{in}} - k_{\text{out}}(1+E)R \\ 0 \end{bmatrix} dt + \begin{bmatrix} \sigma_A & 0 & 0 \\ 0 & \sigma_R & 0 \\ 0 & 0 & \sigma_E \end{bmatrix} d\boldsymbol{w}_t, \qquad (29)
$$

where the initial state of E is denoted by E_0.

The parameter estimates of the second model iteration in (29) are shown in Table I. σ_E is the only diffusion term which is significant while both σ_A and σ_R are insignificant on a 5% significance level. The significant diffusion term σ_E confirms the suspicion that E is deficiently described by a constant.

In order to deconvolve the functional structure of E, the next step in the second grey-box PK/PD modelling iteration is to use the EKF filter estimates together with non-parametric modelling (2). The filter estimates $\hat{A}_{j|j}$, $\hat{R}_{j|j}$, and $\hat{E}_{j|j}$ at each time instant t_j are obtained by using the parameter estimates of the re-formulated model in (29) along with the measurements until time t_j by means of the EKF. The partial dependence plots of $\hat{E}_{j|j}$ vs. $\hat{A}_{j|j}/V$ and $\hat{R}_{j|j}$ are shown in Fig. 2 together with a fitted local regression model. The plots clearly indicate that $\hat{E}_{j|j}$ is highly dependent on $\hat{A}_{j|j}/V$ and the functional relationship seems to be a sigmoidal
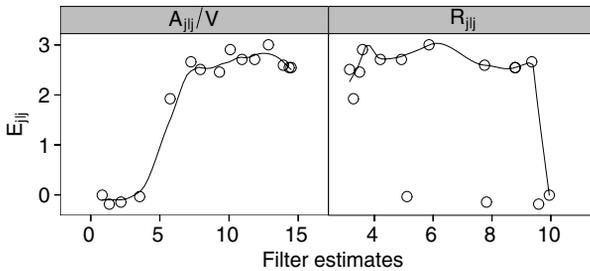

Fig. 2. Partial dependence plots of $\hat{E}_{j|j}$ vs. $\hat{A}_{j|j}/V$ (Left) and $\hat{E}_{j|j}$ vs. $\hat{R}_{j|j}$ (Right). Filter estimates (circles) and fitted local regression model (solid lines).

$E_{\max}$ model. Since $\hat{E}_{j|j}$ does not seem to depend on $\hat{R}_{j|j}$, the dependency of R on E is disregarded in the next model iteration.

The assumption about E being constant should therefore be replaced in the next re-formulated model with an assumption about E being a sigmoidal $E_{\max}$ function of the PK concentration, i.e., A/V. The re-formulated system equations in the third (and last) model iteration in the grey-box PK/PD modelling cycle iteration are therefore

$$\begin{bmatrix} dA \\ dR \end{bmatrix} = \begin{bmatrix} -\frac{CL}{V}A \\ K_{\text{in}} - k_{\text{out}}(1 + \frac{E_{\max}(A/V)^{\gamma}}{EC_{50}^{\gamma}+(A/V)^{\gamma}})R \end{bmatrix} dt + \begin{bmatrix} \sigma_A & 0 \\ 0 & \sigma_R \end{bmatrix} d\boldsymbol{w}_t. \tag{30}$$

The estimated parameters of the third model iteration in (30) are shown in Table I. The estimated system noise parameters are both insignificant on a 5% significance level indicating that the system model is adequate. The remaining PK/PD parameter estimates are all reasonably close to their simulated values. The predicted PK/PD concentration–time profiles are shown in Fig. 3 and model (30) seems to capture the true dynamics of the simulated PK/PD system in (28). The grey-box PK/PD modelling cycle iteration is thus terminated and the grey-box PK/PD modelling framework has shown its usefulness in uncovering the true model structure through iterative model development.

Glucose Minimal Model

The glucose minimal model (GMM) initially proposed by Dr. Richard N. Bergman et al., (4,5) is used to model the clinical data from an IVGTT. The model was initially introduced for the interpretation
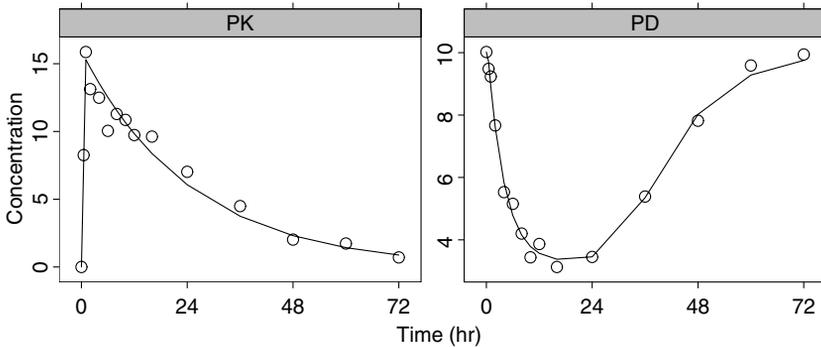

Fig. 3. Plot of observed (circles) and predicted (solid lines) concentrations from the third grey-box model iteration. PK (Left) and PD (Right).

of the plasma glucose and insulin concentrations following an IVGTT in order to assess metabolic indices of subjects likely to develop diabetes.

The GMM consists of two compartments, i.e., a glucose and remote insulin compartment. The controlling action of insulin on hepatic glucose production and utilization of glucose in the peripheral tissue is modelled using the insulin in the remote compartment. The stochastic differential equations governing the GMM are

$$\begin{bmatrix} dG \\ dX \end{bmatrix} = \begin{bmatrix} S_G G_b - (S_G + X)G \\ -p_2(X - S_I(I - I_b)) \end{bmatrix} dt + \begin{bmatrix} \sigma_G & 0 \\ 0 & \sigma_X \end{bmatrix} d\boldsymbol{w}_t, \qquad (31)$$

where G is the state vector for the plasma glucose concentration and X is the insulin action which is proportional to the insulin in the remote compartment. G_b and I_b are the basal plasma glucose and insulin concentrations, respectively, and typically measured as the mean of the glucose measurements before the glucose bolus injection. I is the observed plasma insulin concentration which is used as a forcing function in the GMM. The parameters S_G and S_I are the insulin-independent and insulin-dependent rate constants of glucose uptake commonly referred to as the glucose effectiveness and the insulin sensitivity index, respectively. The parameter p_2 corresponds to the spontaneous decrease of tissue glucose uptake ability (14). The diffusion term in (31) is parameterized as a diagonal matrix since the system noise is assumed independent between the glucose and insulin action states. This will furthermore help pinpoint model deficiencies as in the simulated grey-box PK/PD example. The glucose concentration at $t = 0$ min is estimated as the parameter G_0 while the insulin action at $t = 0$ min per definition is zero.

In order to obtain homogeneity of the measurement error variance, it is modelled using an additive error model on the log-scale corresponding to a constant coefficient of variation (CV) model on the untransformed scale, i.e.,

$$\log G_{\text{obs},j} = \log G_j + e_j, \quad e_j \in N(0, \sigma^2), \qquad (32)$$

where $G_{\text{obs},j}$ and G_j are the observed and predicted glucose concentrations, respectively, at time instants t_j. The measurement error terms e_j are normal distributed with mean zero and variance σ^2. The first 7 min of observed glucose concentrations were set as missing for the parameter estimation because of the single compartment assumption of the glucose distribution (15).

The unknown parameters in (31) are estimated using ML estimation. A standard two-stage approach (16) was used to obtain the two first moments of the population distribution by calculating the empirical mean and variance of all the individual parameter estimates. The empirical mean and relative standard deviation (RSD) estimates of the 244 subjects from the IVGTT are shown in Table II. The grey-box estimates using SDEs are compared to estimates using ODEs, i.e., where the diffusion term σ_w is set to zero in the SDEs. The main difference between the SDE and ODE parameter estimates is in the insulin sensitivity index where the empirical mean and RSD estimates for the SDE and ODE approach are 1.94×10^{-4} (150%) and 3.31×10^{-4} $\min^{-2}$ pM^{-1} (358%), respectively. The estimated measurement error is furthermore deflated from 5.94% using ODEs to 2.88% using SDEs since residual error due to model mis-specification is accounted for in the system noise parameters. The system noise parameter σ_G is significant for almost all subjects while σ_X is insignificant. This indicates that the GMM does not fully capture the dynamics of the IVGTT. The GMM has been reassessed in (17–20) but these models are not considered in this work since they require the injected glucose to be labelled to be able to estimate yet another unobservable state.

The predicted glucose concentration–time profiles using SDEs and ODEs are shown in Fig. 4 along with the observed glucose concentrations for the 244 subjects. The SDE and ODE predictions seem very similar but the drop in the glucose concentrations below the basal glucose level is not so significant in the SDE predictions. The goodness-of-fit graphs of observed versus predicted glucose concentrations are displayed in Fig. 5 for the SDE and ODE approach.

Table II. Empirical mean and relative standard deviation (RSD) estimates for 244 subjects. Comparison of standard two-stage GMM parameter estimates using stochastic (SDE) and ordinary (ODE) differential equations.

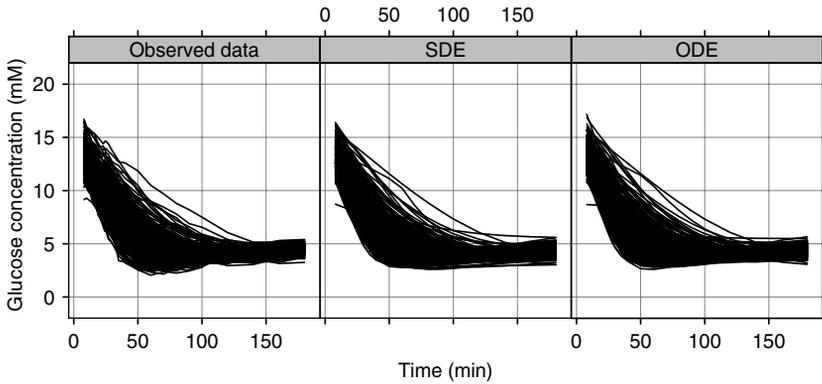| Method		SDE		ODE	
Parameter	Unit	Mean	RSD (%)	Mean	RSD (%)
G_0	[mM]	14.7	10.2	14.9	10.0
S_G	[$\min^{-1}$]	0.0181	35.7	0.0170	34.0
S_I	[$\min^{-2}$ pM^{-1}]	0.000194	150	0.000331	356
P_2	[$\min^{-1}$]	0.0336	183	0.0324	118
σ_G	[mM $\min^{-1/2}$]	0.0457	90.8	–	–
σ_X	[$\min^{-3/2}$]	0.000196	271	–	–
σ	[%]	2.88	197	5.94	127

Fig. 4. Plot of observed and predicted glucose concentrations with each line representing data from one individual. Observed data (Left), SDE predictions (Middle), and ODE predictions (Right).
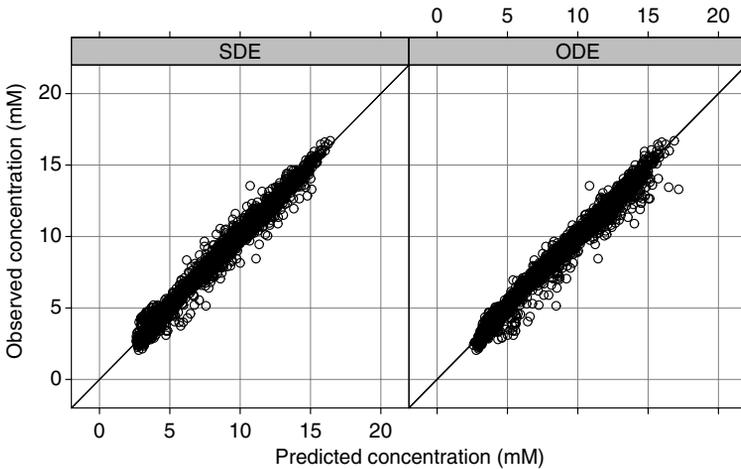

Fig. 5. Plot of observed glucose concentrations vs. SDE (Left) and ODE (Right) predictions. The solid lines are the lines of identity.

The estimates of S_G from the GMM are generally less precise than those of S_I. This is partially because of the model assumption concerning the self-production of glucose which is assumed to be negligible. The glucose effectiveness S_G does not only reflect the glucose effect on glucose uptake and production but also the rapid exchange of glucose between accessible and non-accessible glucose pools in the early part of the IVGTT

according to Caumo *et al.* (18). As mentioned previously, the system noise parameter σ_G is significant for almost all subjects in the IVGTT which indicates a potential model deficiency. The assumption about S_G being constant is therefore examined by including S_G as an additional state variable with the drift term set equal to zero since it is considered to be a constant plus a diffusion parameter σ_{S_G}, i.e.,

$$\begin{bmatrix} dG \\ dX \\ dS_G \end{bmatrix} = \begin{bmatrix} S_G G_b - (S_G + X)G \\ -p_2(X - S_I(I - I_b)) \\ 0 \end{bmatrix} dt + \begin{bmatrix} \sigma_G & 0 & 0 \\ 0 & \sigma_X & 0 \\ 0 & 0 & \sigma_{S_G} \end{bmatrix} d\boldsymbol{w}_t. \qquad (33)$$

The parameter estimates of the diffusion term σ_{S_G} in the system equation for S_G is significant while the diffusion term parameters σ_G and σ_X now both are insignificant on a 5% significance level (not shown). The partial dependency plots of $\hat{S}_{G,j|j}$ vs. $\hat{X}_{j|j}$ and $\hat{G}_{j|j}$ for a representative subject from the study are shown in Fig. 6.

There does not seem to be a relationship between $\hat{S}_{G,j|j}$ and $\hat{X}_{j|j}$ in Fig. 6. $\hat{S}_{G,j|j}$ seems to depend more strongly on $\hat{G}_{j|j}$ where $\hat{S}_{G,j|j}$ assume values around 0.06 min^{-1} at glucose concentrations around 5 mM while $\hat{S}_{G,j|j}$ is around 0.02 min^{-1} at glucose concentrations above 8 mM. The observed pattern in the partial dependency plot of $\hat{S}_{G,j|j}$ vs. $\hat{G}_{j|j}$ in Fig. 6 might be a result of S_G being a composite parameter reflecting the inability of the single-compartment approximation in the GMM. The functional relationship between $\hat{S}_{G,j|j}$ and $\hat{G}_{j|j}$ is however not easily seen in Fig. 6 and the next step in the grey-box PK/PD modelling cycle iteration is not obvious. However, it may be concluded that the simplifying assumption about S_G being constant seems to be wrong.
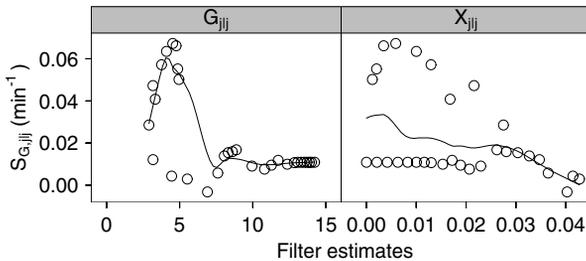

Fig. 6. Partial dependence plots of $\hat{S}_{G,j|j}$ vs. $\hat{G}_{j|j}$ (left) and $\hat{S}_{G,j|j}$ vs. $\hat{X}_{j|j}$ (right) for a representative subject.

CONCLUSION

Simulated data of a one-compartment PK model with IV infusion and an indirect response PD model with stimulation of k_{out} has successfully been used to illustrate the methodology behind the grey-box PK/PD modelling framework for uncovering model deficiencies through iterative model development. It was furthermore shown how to use the grey-box PK/PD modelling framework for deconvolution of the functional relationship between PK and PD data.

The glucose minimal model was used to model clinical data from an IVGTT where the obtained estimate of the insulin sensitivity index S_I was significantly different between the SDE and ODE approach. The estimated diffusion parameters σ_G and σ_X indicated that the GMM is too simple and not fully adequate for describing the complicated dynamic *in vivo* system of insulin and glucose following an IVGTT. In the attempt to test the assumption about the glucose effectiveness S_G being constant, S_G was included as another state variable with the drift term set equal to zero plus a diffusion term. The partial dependence plot of $\hat{S}_{G,j|j}$ vs. $\hat{G}_{j|j}$ clearly showed that the assumption about S_G being constant is not valid.

ACKNOWLEDGEMENTS

REFERENCES

1. H. Madsen and J. Holst. Modelling Non-linear and Non-stationary Time Series. *Informatics and Mathematical Modelling (IMM). DTU*, 2000.
2. N. R. Kristensen, H. Madsen, and S. B. Jørgensen. A method for systematic improvement of stochastic grey-box models. *Comput. Chem. Eng.* **28** (8):1431–1449 (2004).
3. W. J. Jusko and H. C. Ko. Physiologic indirect response models characterize diverse types of pharmacodynamic effects. *Clin. Pharmacol. Ther.* **56** (4):406–419 (1994).
4. R. N. Bergman, Y. Z. Ider, C. R. Bowden, and C. Cobelli. Quantitative estimation of insulin sensitivity. *Am. J. Physiol. Endocrinol. Metab.* **236** (6):E667–E677 (1979).
5. R. N. Bergman, L. S. Phillips, and C. Cobelli. Physiologic Evaluation of Factors Controlling Glucose Tolerance in Man. *J. Clin. Invest.* **68:**1456–1467 (1981).
6. A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, New York, (1970).
7. C. W. Tornøe, J. L. Jacobsen and H. Madsen. Grey-box pharmacokinetic/pharmacodynamic modelling of a Euglycaemic Clamp Study. *J. Math. Biol.* **48:**591–604 (2004).
8. C. W. Tornøe. *Grey-box PK/PD Modelling of Insulin*. Master's thesis. IMM, DTU (2002).
9. N. R. Kristensen, H. Melgaard and H. Madsen. CTSM 2.3, User's and Mathematics Guide, http://www.imm.dtu.dk/CTSM (2003).

10. B. Øksendal. *Stochastic Differential Equations – An Introduction with Applications, 5. edn.* Springer-Verlag, Berlin, Germany (1998).

11. D. Z. D. Argenio and K. Park. Uncertain pharmocokinetic/pharmocodynamic systems: design, estimation and control. *Control Eng. Pract.* **5** (12):1707–1716 (1997).

12. M. Ramanathan. An application of Ito's lemma in population pharmacokinetics and pharmacodynamics. *Pharm. Res.* **16** (4):584–586 (1999).

13. M. Ramanathan. Pharmacokinetic variability and therapeutic drug monitoring actions at steady state. *Pharm. Res.* **17** (5):589–592 (2000).

14. A. De Gaetano and O. Arino. Mathematical modelling of the Intravenous Glucose Tolerance test. *J. Math. Biol.* **40**:136–168 (2000).

15. G. Pacini and R. N. Bergman. MINMOD: A computer program to calculate insulin sensitivity and pancreatic responsivity from the frequently sampled intravenous glucose tolerance test. *Comput. Meth. Prog. Biomed.* **23**:113–122 (1986).

16. M. Davidian and D. M. Giltinan. *Nonlinear Models for Repeated Measurement Data.* Chapman & Hall, London (1995).

17. A. Caumo and C. Cobelli. Hepatic glucose production during the labelled IVGTT: estimation by deconvolution with a new minimal model. *Am. J. Physiol. Endocrinol. Metab.* **264**:E829–E841 (1993).

18. A. Caumo, P. Vicini, J. J. Zachwieja, A. Avogaro, K. Yarasheski, D. M. Bier, and C. Cobelli. Undermodelling affects Minimal Model Indexes: insights from a two-compartment model. *Am. J. Physiol. Endocrinol. Metab.* **276**:E1171–E1193 (1999).

19. A. Mari. Assessment of insulin sensitivity and secretion with the labelled intravenous glucose tolerance test: Improved modelling analysis. *Diabetologia* **41**:1029–1039 (1998).

20. T.-C. Ni, M. Ader, and R. N. Bergman. Reassessment of glucose effectiveness and insulin sensitivity from minimal model analysis: A theoretical evaluation of the single-compartment glucose distribution assumption. *Diabetes* **46**:1813–1821 (1997).