PK/PD Modelling using Stochastic Differential Equations

FMS and DSBS Workshop

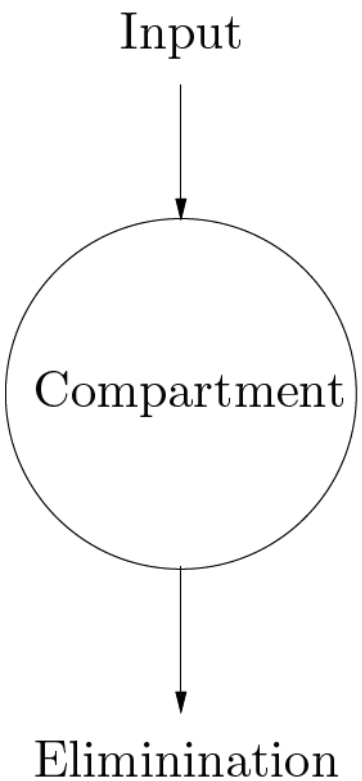
Henrik Madsen

DTU Informatics

hm@imm.dtu.dk

Outline

Problem Scenario

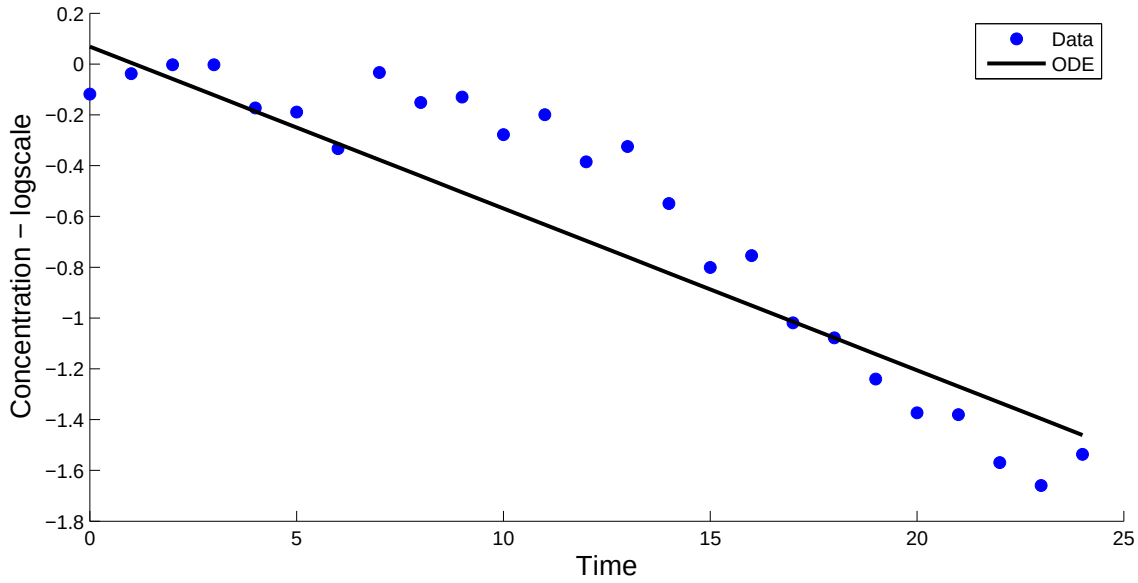


Ordinary differential equation

$$dA = -KA\,dt$$

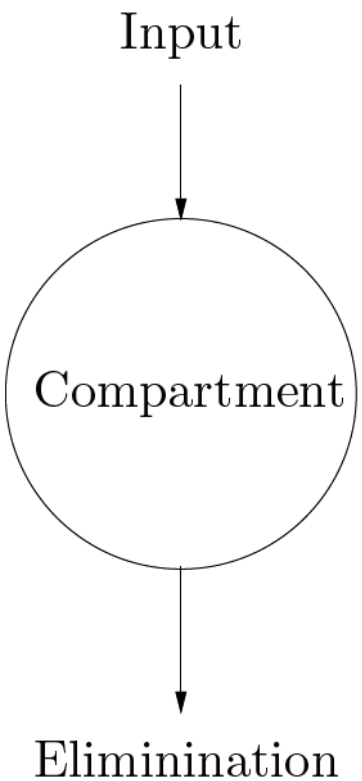
$$Y = A + \epsilon$$

ODE



- Correlated residuals!!

Problem Scenario

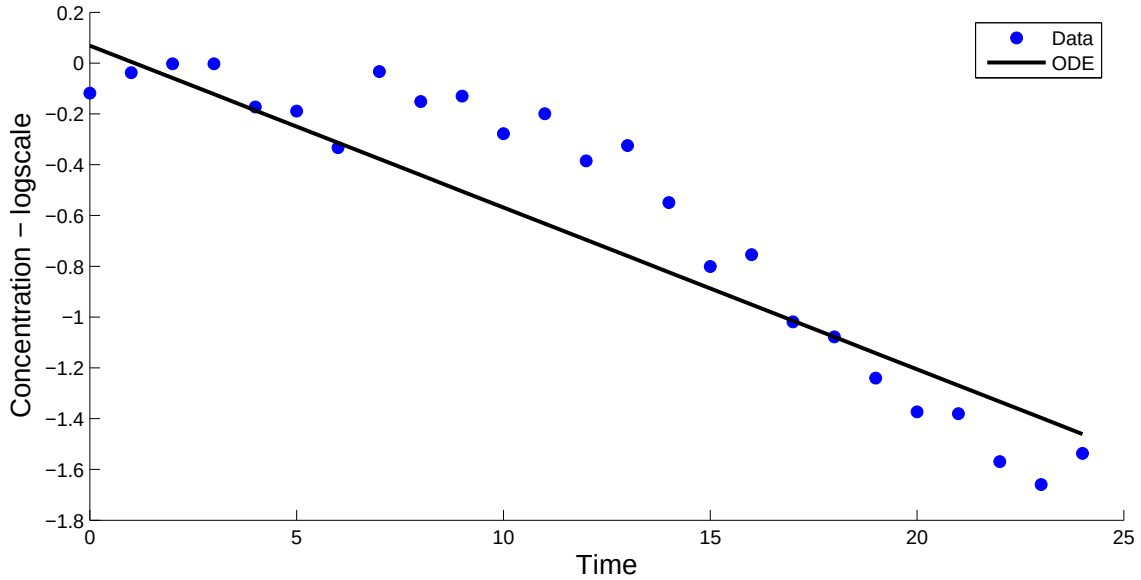


Stochastic differential equation

$$dA = -KA\,dt + dw$$

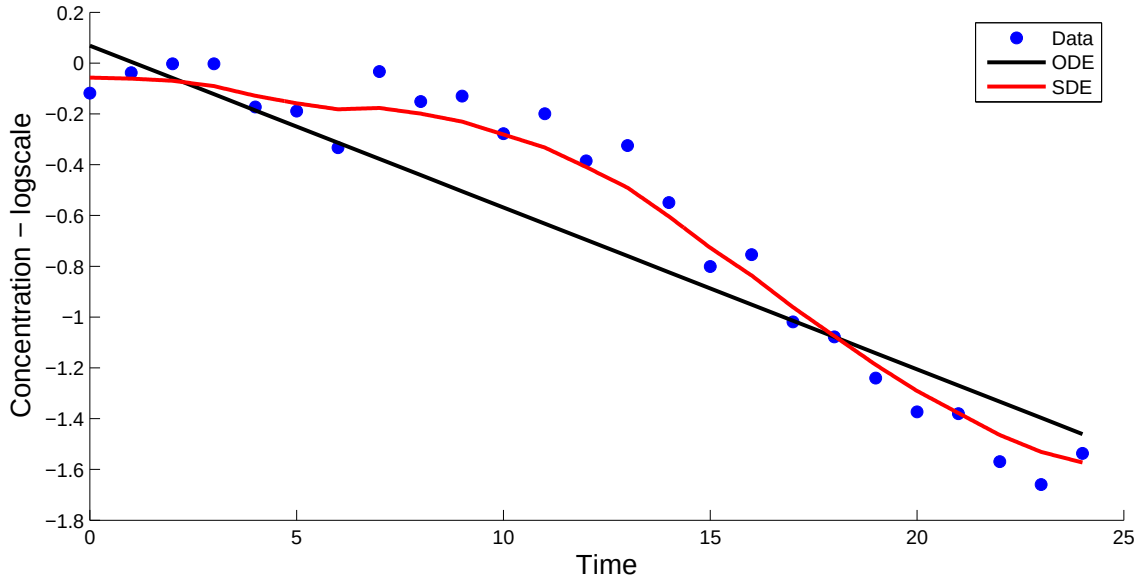
$$Y = A + e$$

ODE vs SDE



- Correlated residuals
- System noise
- Measurement noise

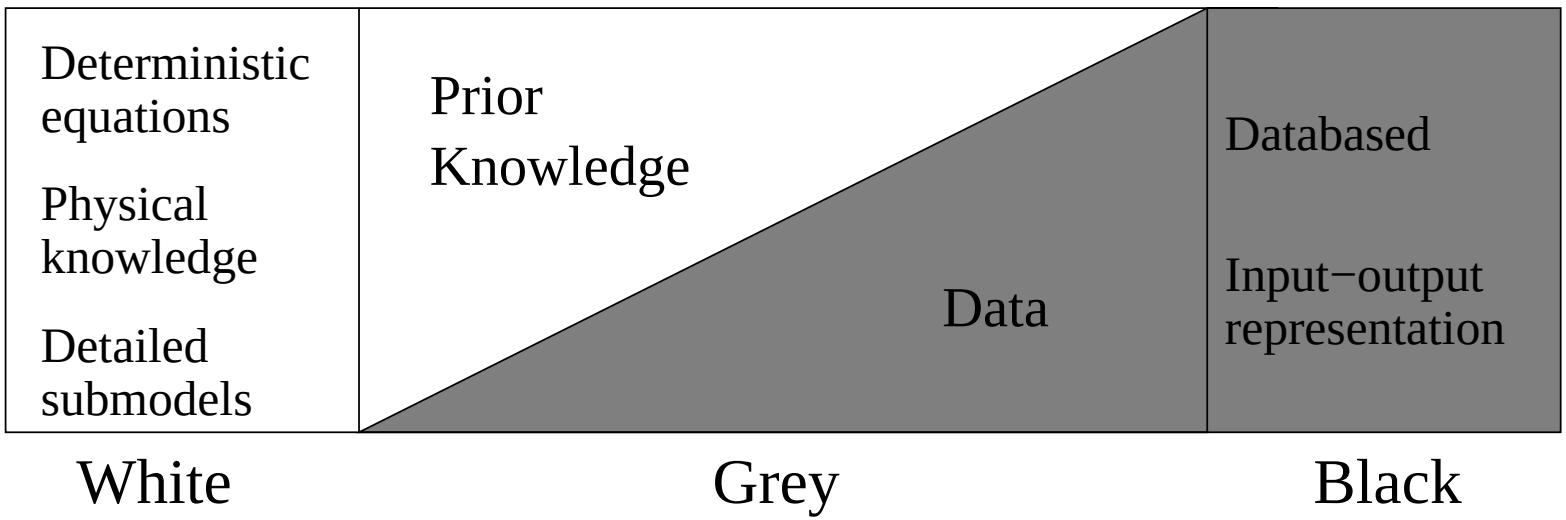
ODE vs SDE



- Correlated residuals
- System noise
- Measurement noise

The grey box modelling concept

- Combines prior physical knowledge with information in data.
- The model is not completely described by physical equations, but equations and the parameters are physically interpretable.

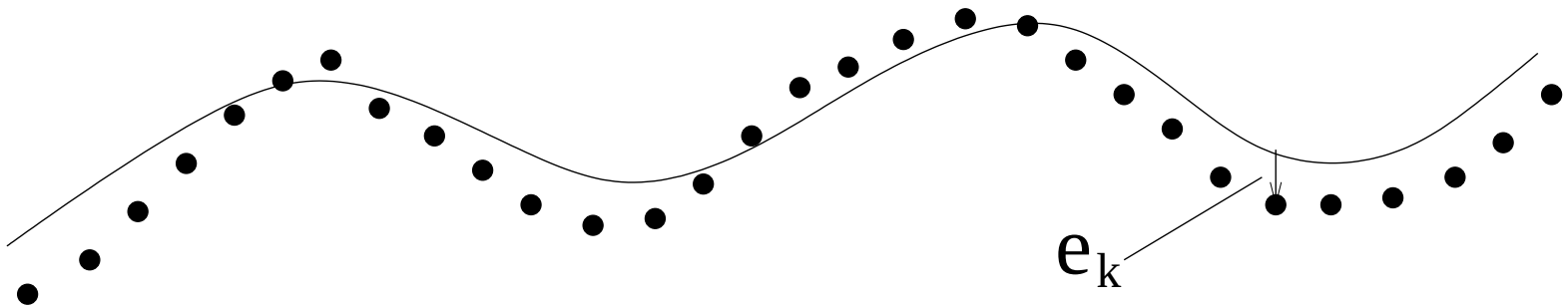


The box in the middle (denoted Grey) may be expressed as the relationship between the amount of prior knowledge and data available.

Why use grey box modelling?

- Prior physical knowledge can be used.
- Non-linear and non-stationary models are easily formulated.
- Missing data are easily accommodated.
- It is possible to estimate state variables that are not measured.
- Available physical knowledge and statistical modelling tools is combined to estimate the parameters of a rather complex dynamical system.
- The parameters contain information from the data that can be directly interpreted by the scientist.
- Fewer parameters $\rightarrow$ more power in the statistical tests.
- The physical expert and the statistician can collaborate in the model formulation.

Stochastic Differential Equations (SDE's)



- The line demonstrates a model prediction, whereas the dots denote typical observations.
- Notice: Autocorrelated residuals are most often seen
  - calls for using Stochastic Differential Equations (SDE's) as an alternative to Ordinary Differential Equations (ODE's).

The continuous-discrete time non-linear stochastic state space model – The system equation (a set of Itô stochastic differential eqs.)

The system equation consists of a *drift* and a *diffusion* term.

$$\mathrm{d}\boldsymbol{X}_t = f(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{\theta})\,\mathrm{d}t + G(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{\theta})\,\mathrm{d}\boldsymbol{W}_t, \quad \boldsymbol{X}_{t_0} = \boldsymbol{X}_0$$

Notation

$\boldsymbol{X}_t \in \mathbb{R}^n$	State vector
$\boldsymbol{U}_t \in \mathbb{R}^r$	Known input vector
f	Drift term
G	Diffusion term
$\boldsymbol{W}_t$	Wiener process of dimension, d, with incremental covariance $\boldsymbol{Q}_t$
$\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$	Unknown parameter vector

The observation equation

The observations are in discrete time, functions of state, input, and parameters, and are subject to noise:

$$\boldsymbol{Y}_{t_i} = h(\boldsymbol{X}_{t_i}, \boldsymbol{U}_{t_i}, \boldsymbol{\theta}) + \boldsymbol{e}_{t_i}$$

Notation

$\boldsymbol{Y}_{t_i} \in \mathbb{R}^m$	Observation vector
h	Observation function
$\boldsymbol{e}_{t_i} \in \mathbb{R}^m$	Gaussian white noise with covariance $\boldsymbol{\Sigma}_{t_i}$

Observations are available at the time points t_i : $t_1 < \ldots < t_i < \ldots < t_N$

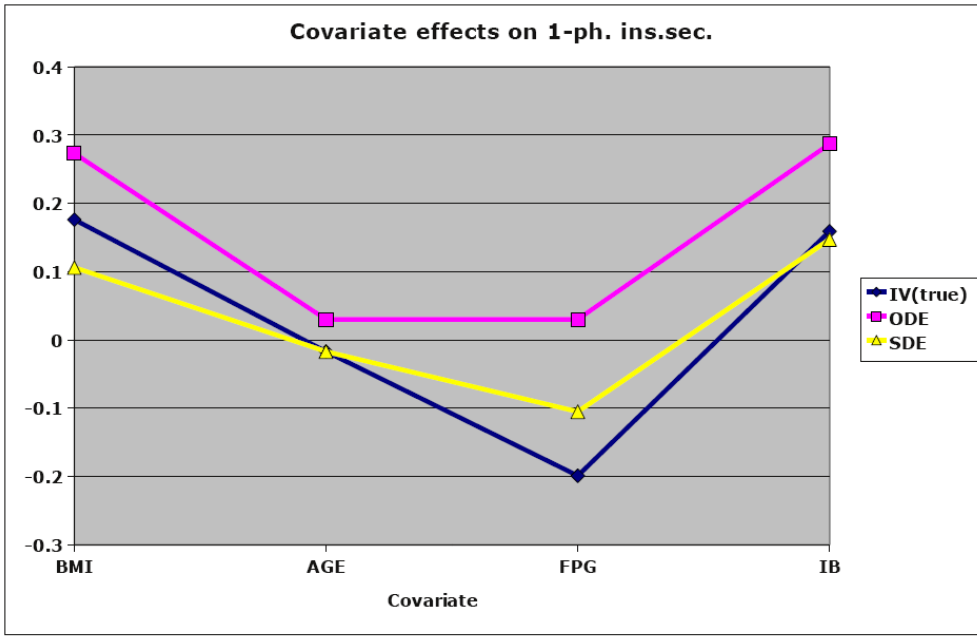
$\boldsymbol{X}_0, \boldsymbol{W}_t, \boldsymbol{e}_{t_i}$ assumed independent for all $(t, t_i), t \neq t_i$

Advantages of using SDE's

- Provides a decomposition of the total error into process error and measurement error.
- Facilitates use of statistical tools for model validation.
- Provides a systematic framework for pinpointing model deficiencies – will be demonstrated later on.
- Covariances of system error and measurement error are estimated.
- SDE based estimation gives more accurate and reliable parameter estimates than ODE based estimation.
- SDEs give more correct (more accurate and realistic) predictions and simulations.
- SDEs give more correct dose (see the following example)

Example: Glucose/Insulin system

- Model predicting the insulin concentration having glucose as input
- **Wrong error model gives wrong dose!**
- SDEs give more reliable estimates of covariate effects



J.B. Møller et.al.: *Predictive performance for population models using stochastic differential equations applied on data from an oral glucose tolerance test*, Journal of PK/PD, 2009

Methods for Identification, Estimation and Model Validation

- **Model Identification:** See the next slide.
- **Parameter Estimation:**
  - (Approx.) Maximum Likelihood Methods
  - Estimation Functions
  - Prediction Error Methods
- **Model Validation:**
  - Test whether the estimated model describes the data.
  - Autocorrelation functions – or Lag Dependent Functions.
  - Other classical methods ...

Identification of Model Structure

- The **diffusion term** gives information for pinpointing model deficiencies.
- Assume that we based on 'large' values of relevant diffusion term(s) suspect $r \in \boldsymbol{\theta}$ to be a function of the states, input or time.
- Then consider the **extended state space model**:

$$\begin{aligned} \mathrm{d}\boldsymbol{X}_t &= f(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{\theta})\,\mathrm{d}t + G(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{\theta})\,\mathrm{d}\boldsymbol{W}_t, \quad \boldsymbol{X}_{t_0} = \boldsymbol{X}_0 \\ \mathrm{d}r_t &= \mathrm{d}W_t^* \\ \boldsymbol{Y}_{t_i} &= h(\boldsymbol{X}_{t_i}, \boldsymbol{U}_{t_i}, \boldsymbol{\theta}) + \boldsymbol{e}_{t_i} \end{aligned} \tag{1}$$

which corresponds to a **random walk** description of r_t.

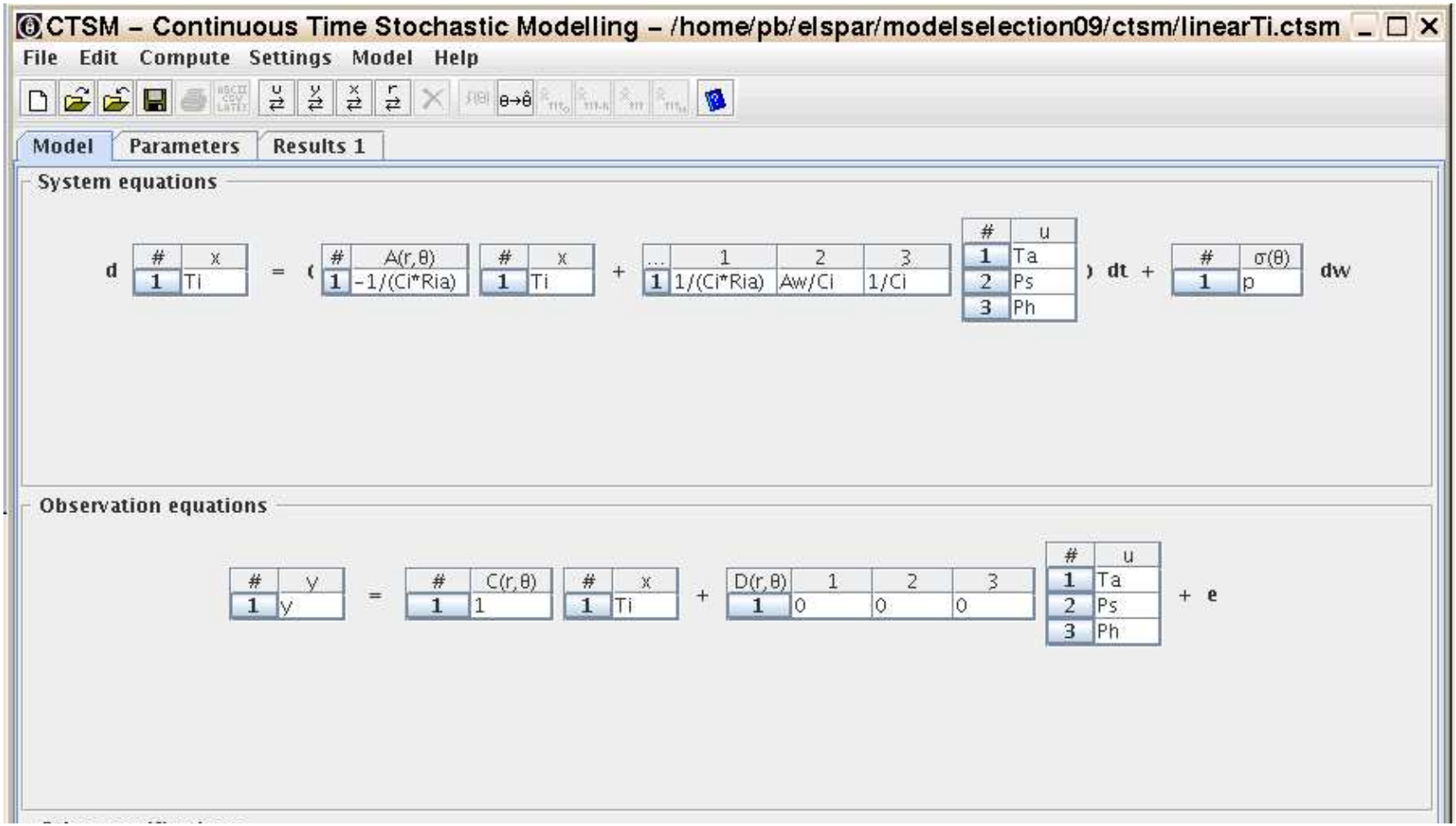
Identification of Model Structure

- Do we observe a significiant reduction of the relevant diffusion term(s)?
- In that case calculate the smoothed state estimate $\widehat{r}_{t|N}$ (use for instance the software tool CTSM - see the next slide).
- Plot $\widehat{r}_{t|N}$ versus the states, inputs and time.
- Identify a possible functional relationship.
- Build that functional relationship into the stochastic state space model.
- Estimate the model parameters and evaluate the improvement – using e.g. likelihood ratio tests.

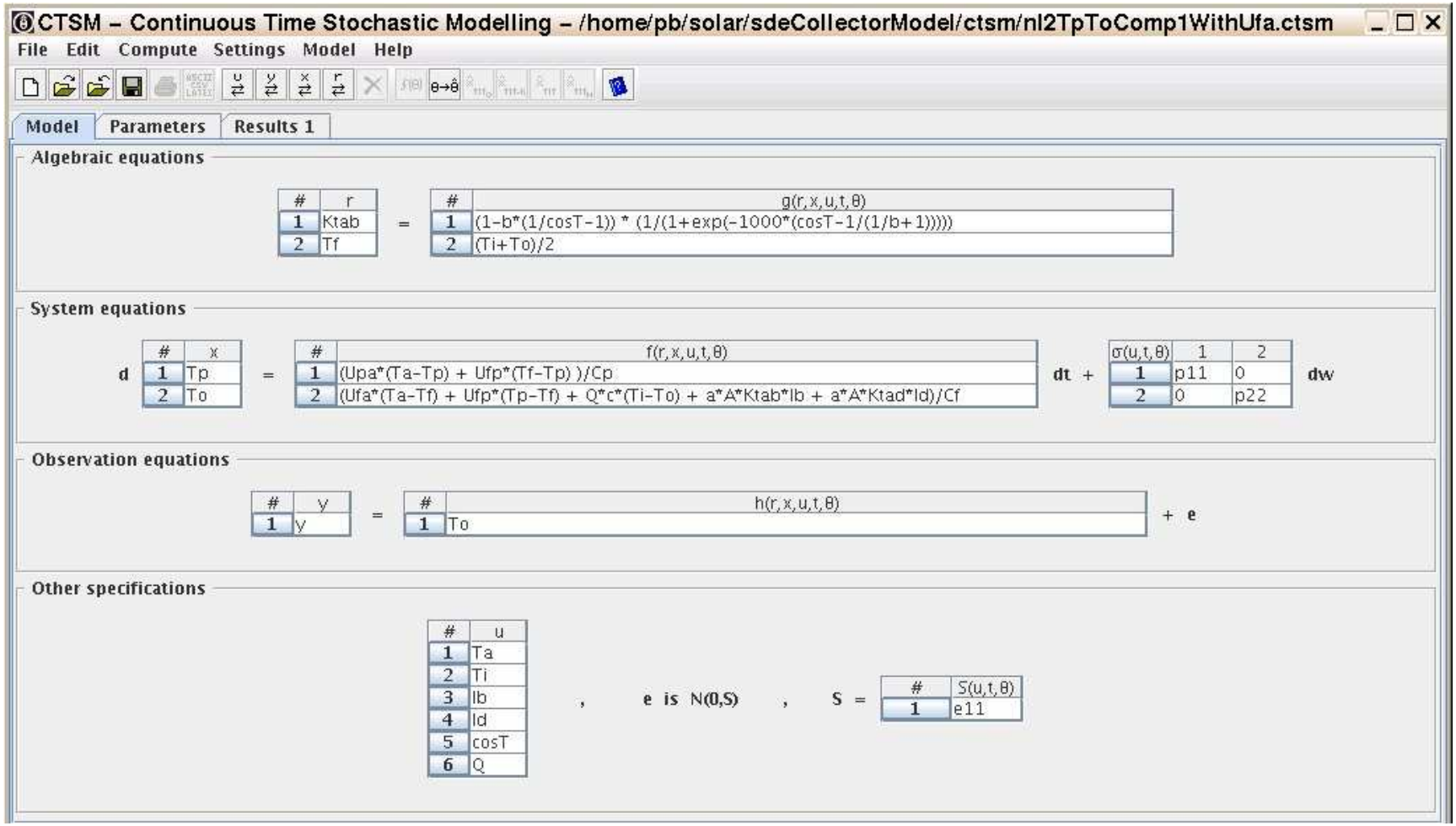
Continuous Time Stochastic Modelling (CTSM)

- The parameter estimation is performed by using the software CTSM.
- The software has been developed at IMM.
- Download from: `www.imm.dtu.dk/ctsm`
- The program analyses the model equations to determine the symbolic names of the parameters to fix and which to estimate.
- The program returns the uncertainty of the parameter estimates as well.

Continuous Time Stochastic Modelling (CTSM) – Linear case



Continuous Time Stochastic Modelling (CTSM) – Non-linear case



Population Modelling – Introduction

- Data originating from several population members/subjects
- Identical experiments
- More data - better estimates of parameters and uncertainties.
- Software: Population Stochastic Modelling (PSM)
- Software download: `www.imm.dtu.dk/psm`

Population Modelling – Introduction

- Data originating from several population members/subjects
- Identical experiments
- More data - better estimates of parameters and uncertainties.
- Software: Population Stochastic Modelling (PSM)
- Software download: `www.imm.dtu.dk/psm`

Nonlinear mixed effects model with SDEs

Population Modelling – Stages

A population model consists of 2 stages

- 1^{st} stage models the process variation within a single population member/subject
- 2^{nd} stage models the variation in parameters between population members/subjects like:

$$\boldsymbol{\phi}_i = g(\boldsymbol{\theta}, \boldsymbol{Z}_i) \cdot exp(\boldsymbol{\eta}_i)$$

$$\boldsymbol{\eta}_i \in N(\mathbf{0}, \boldsymbol{\Omega})$$

Population – Parameter estimation

Parameter estimation using likelihood theory

- Single member/subject - likelihood based on product of conditional densities for each time series of length n_i (called p_1 below).
- Population likelihood is a combination of the random effects η and the single member likelihoods

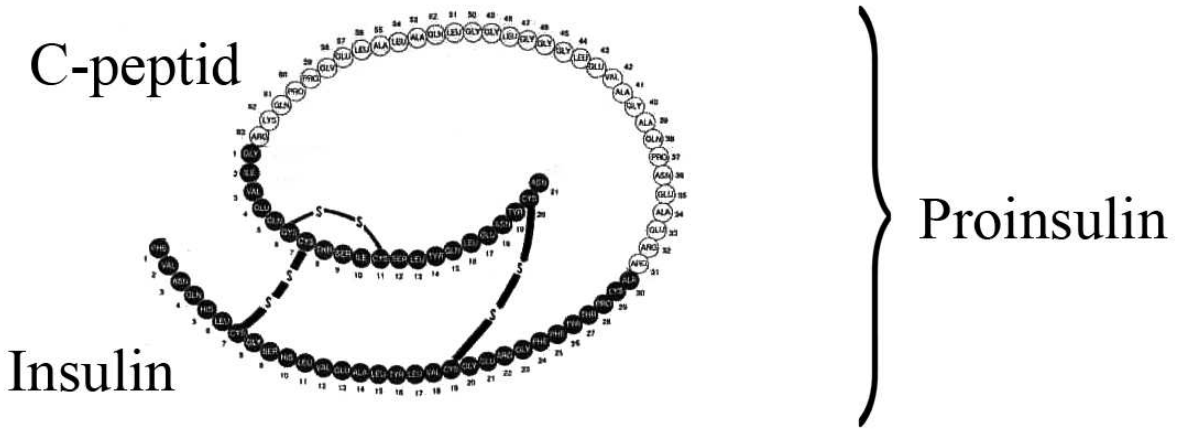
$$L(\boldsymbol{\theta}|\mathcal{Y}_{Nn_i}) \quad \propto \quad \prod_{i=1}^{N} \int p_1(\mathcal{Y}_{in_i}|\boldsymbol{\phi}_i) p_2(\boldsymbol{\eta}_i|\boldsymbol{\Omega}) d\boldsymbol{\eta}_i$$

Diabetes in figures

- 150.000 diagnosed with diabetes in Denmark
  - Treatment costs 2.5B kr./year
  - 150.000 unaware of their diabetes condition
- 171 millions diagnosed world wide
  - Expected to reach 366 million by 2030
- Increased risk for heart diseases, blindness, nerve damage and kidney damage

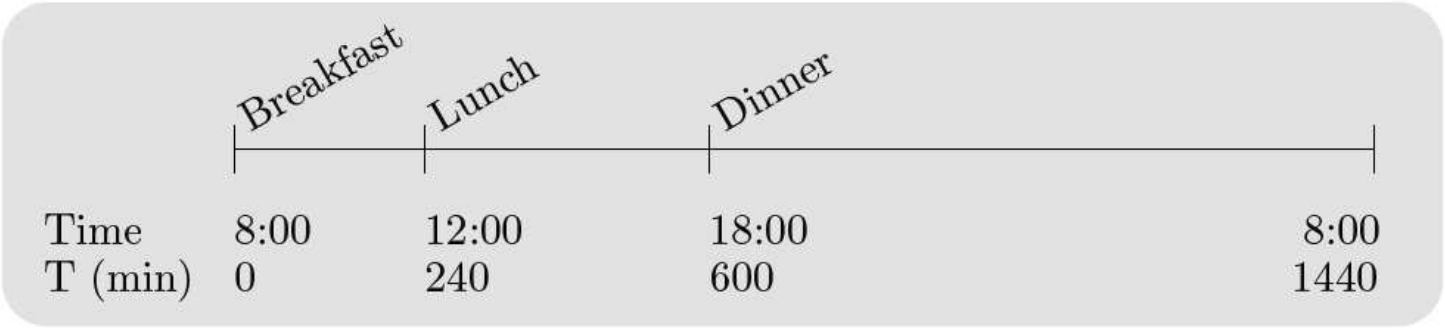
Diabetes physiology

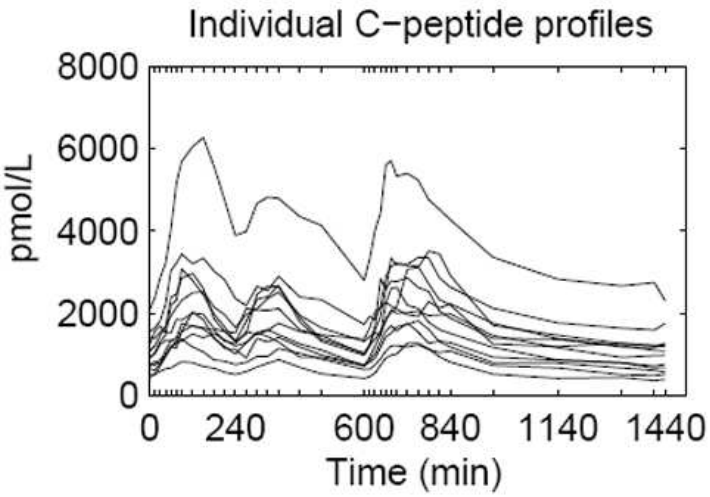
- Insulin is secreted from the Pancreas
  - Extracted by the liver
  - Half-life approx. 5 min.
- C-peptide is co-secreted in equimolar amounts
  - *Not* extracted by the liver
  - Half-life approx. 30 min.

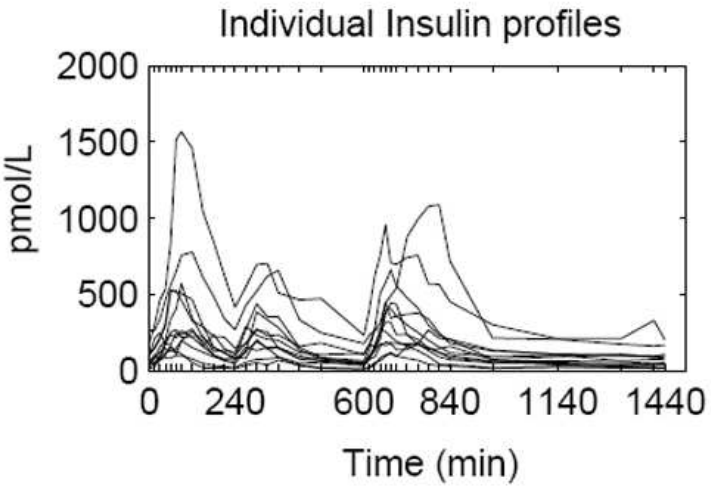


Data – 24H study

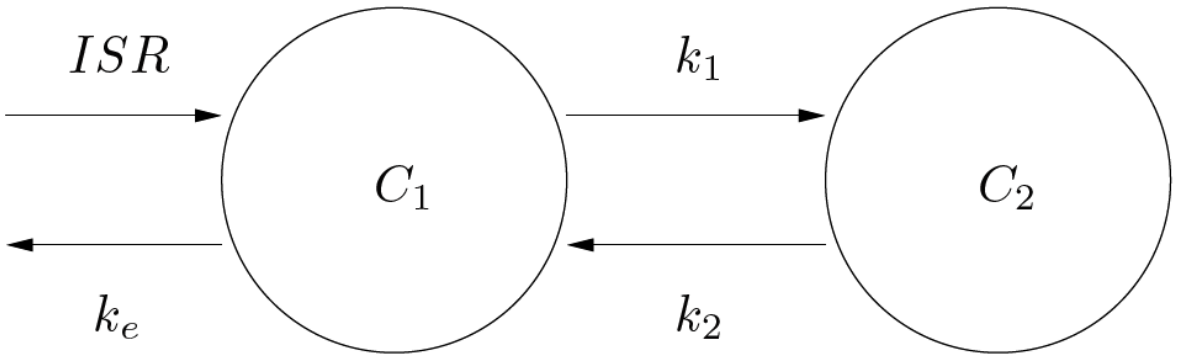
- 12 type 2 diabetic patients
- Three standardized meals







ISR estimate by deconvolution



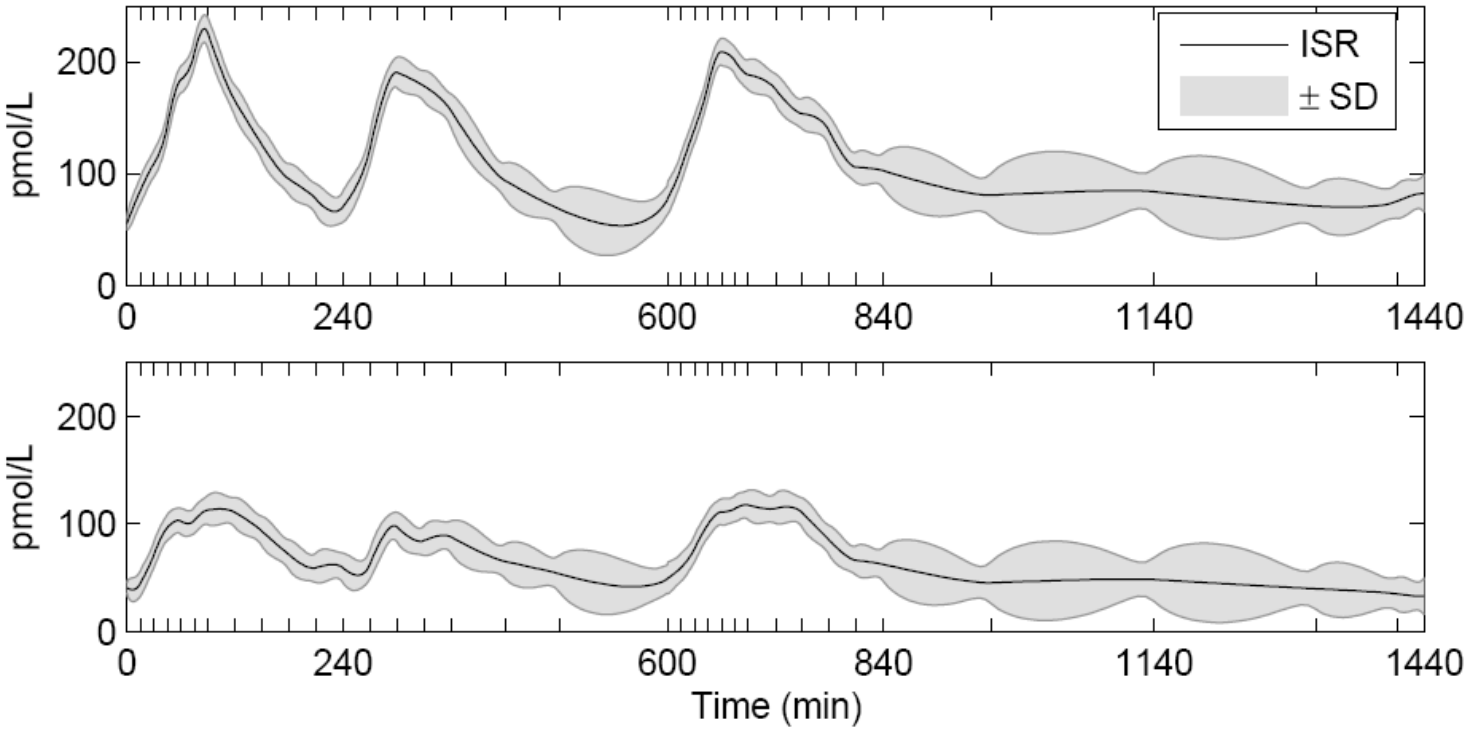
- C-peptide (the observation) is modelled with a 2-compartment model
- ISR modelled as a random walk (the third state in $\boldsymbol{x}$)

$$d\boldsymbol{x} = \begin{bmatrix} -(k_1+k_e) & k_2 & 1 \\ k_1 & -k_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \boldsymbol{x}\, dt + diag \begin{bmatrix} 0 \\ 0 \\ \sigma_{ISR} \end{bmatrix} d\boldsymbol{\omega}$$

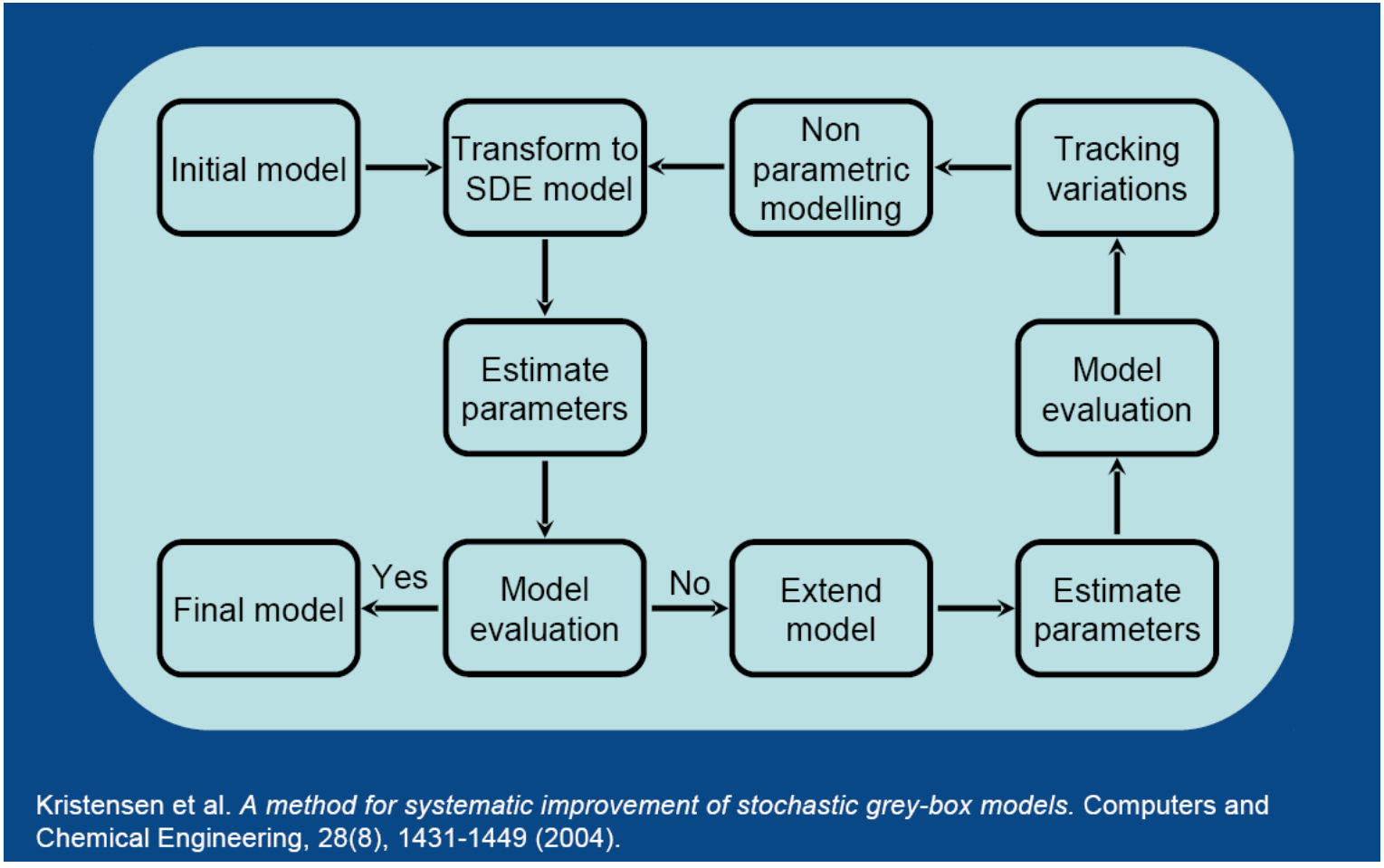
$$y = C_1 + \epsilon$$

ISR estimate by deconvolution

Smoothed estimate of ISR for individual 1 and 2.



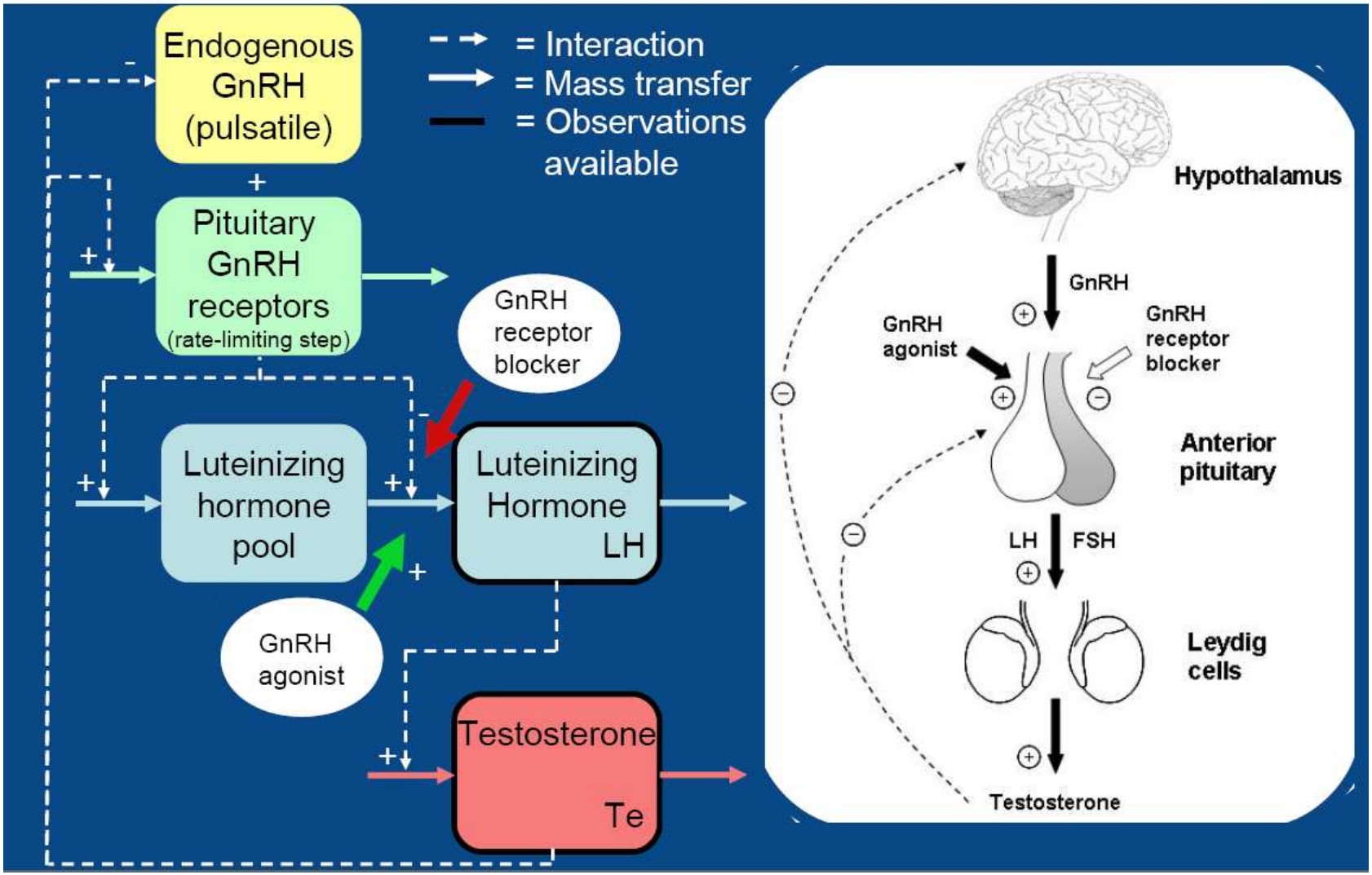
Systematic Modelling Framework



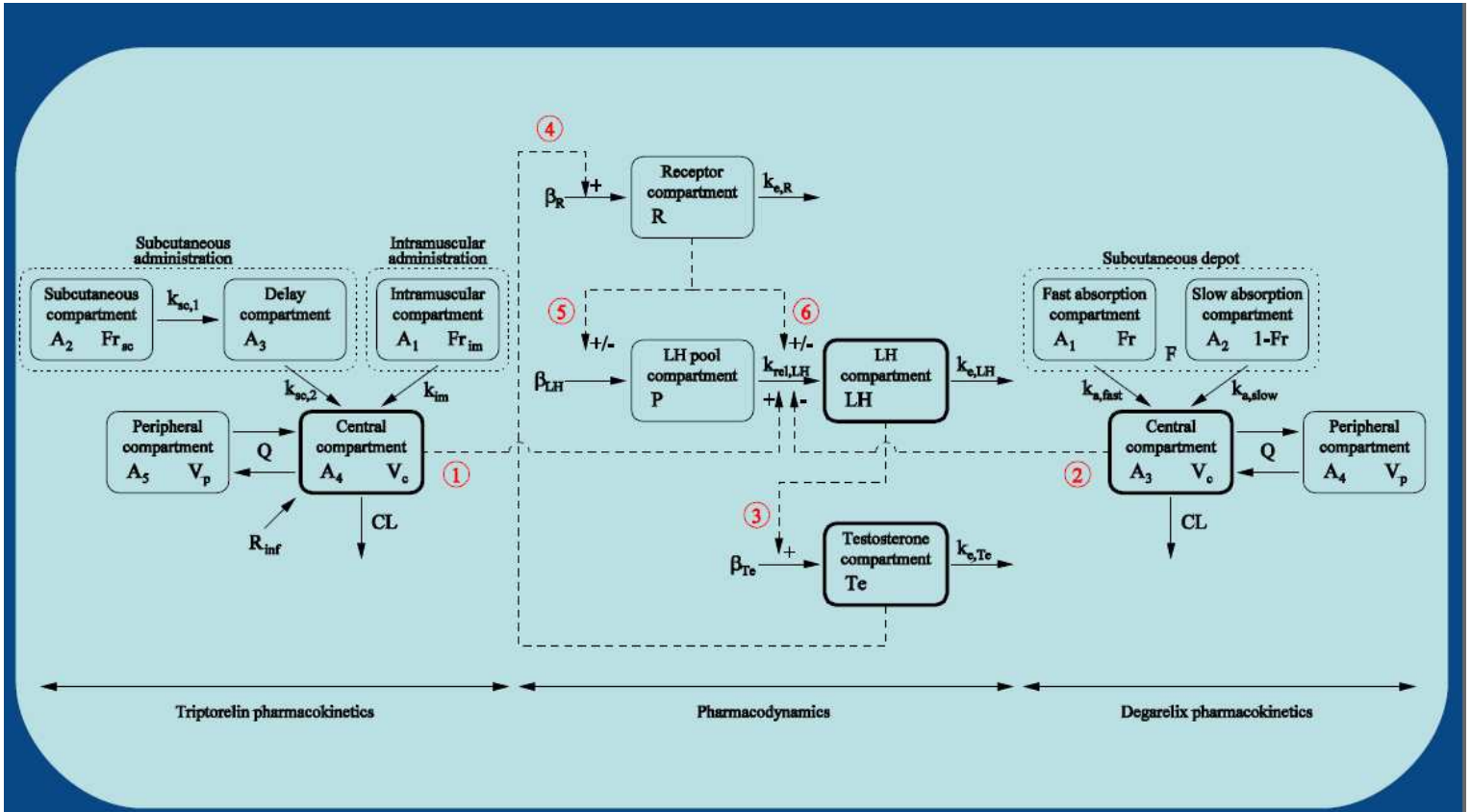
Kristensen et al. *A method for systematic improvement of stochastic grey-box models.* Computers and Chemical Engineering, 28(8), 1431-1449 (2004).

FERRING PHARMACEUTICALS

Mechanism-based model of the HPG axis

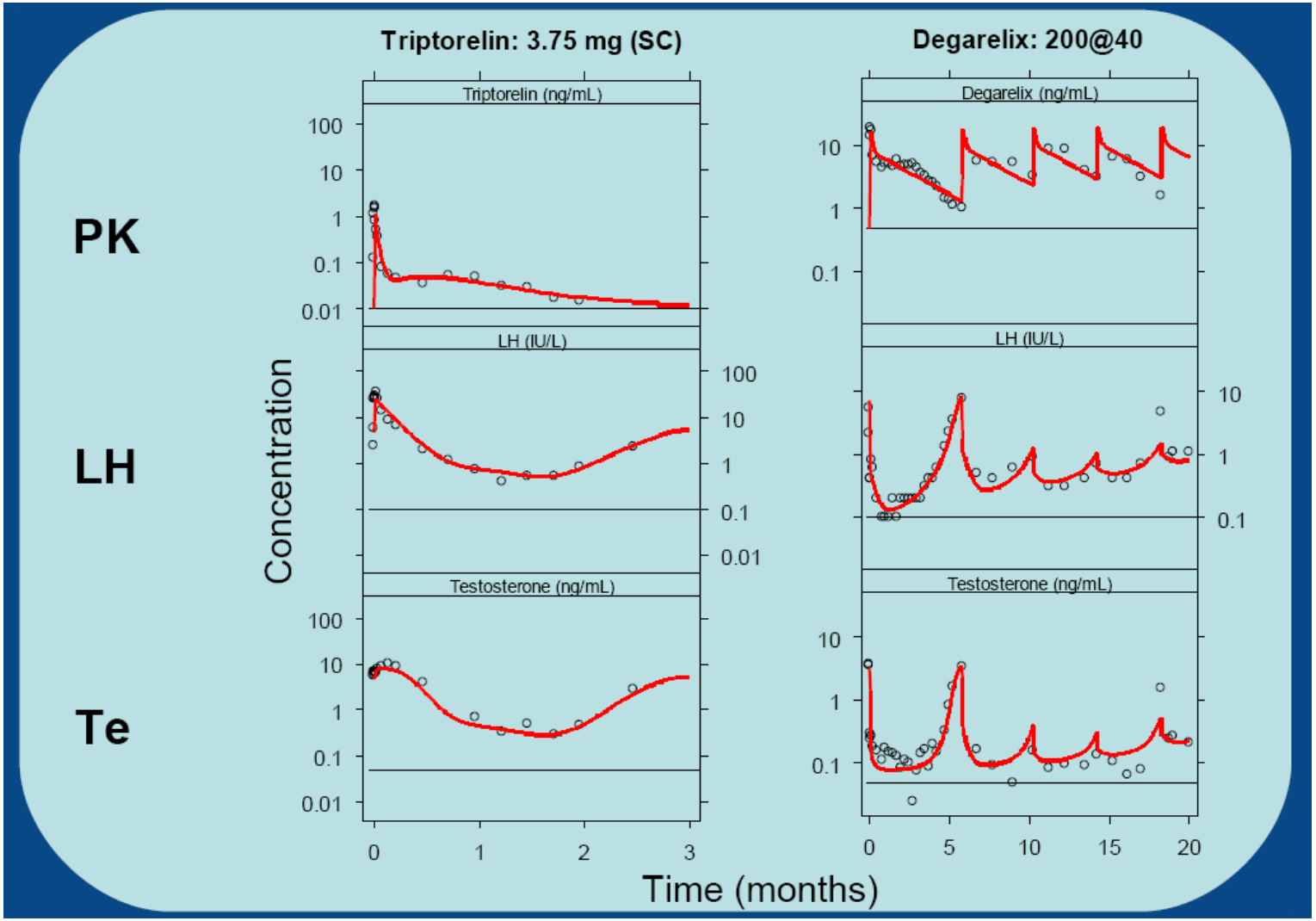


Final model

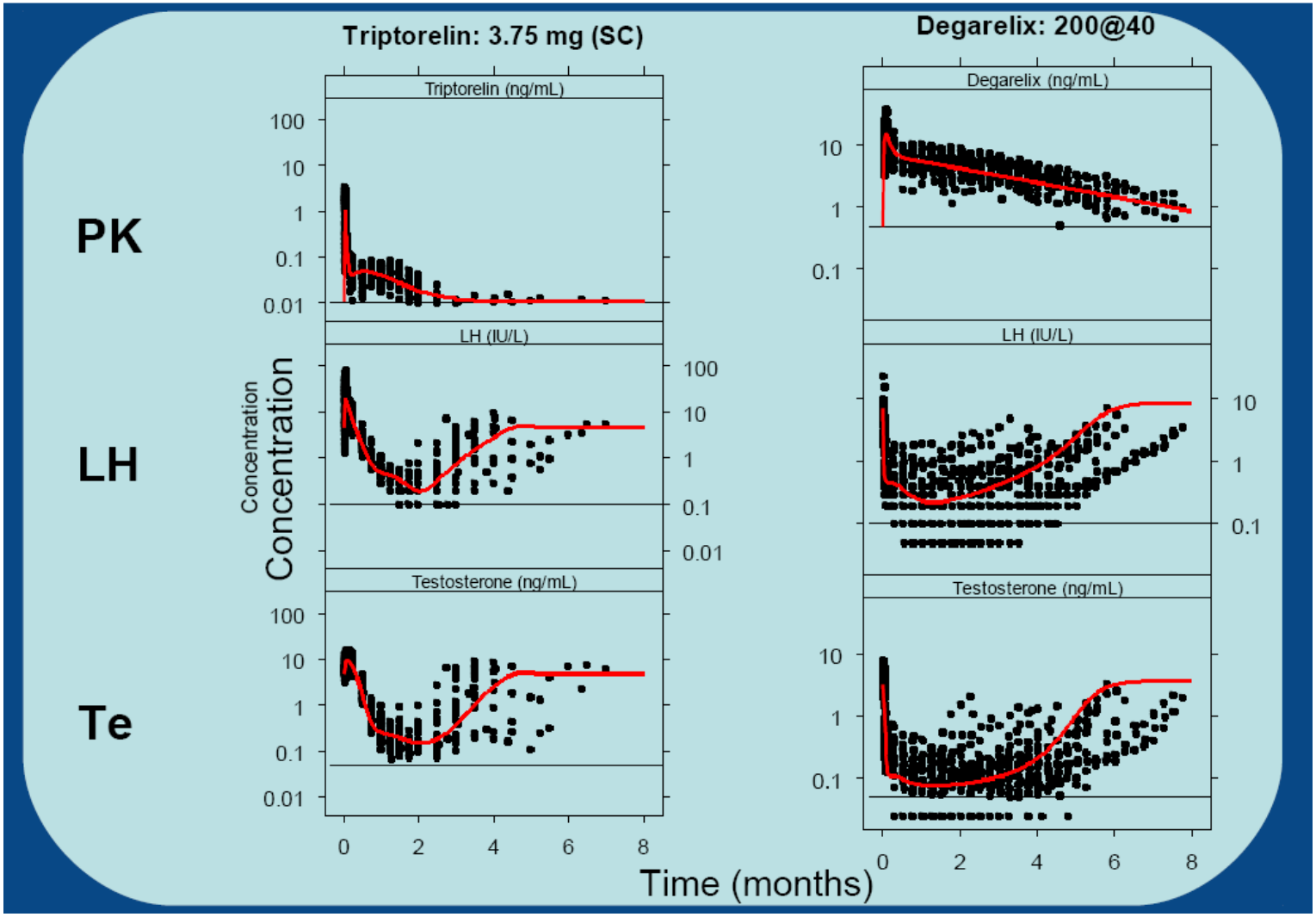


Tornoe et al. *Population PK/PD modelling of the hypothalamic-pituitary-gonadal axis after treatment with GnRH agonist triptorelin and GnRH receptor blocker degarelix*. Submitted to BJCP 2006.

Individual profile



Population profile



Summary

By using stochastic differential equations for PK/PD modelling we have

- better predictions and simulations
- systematic methods for model development
- methods for finding the best model (LR-tests, etc.)
- statistical methods for model validation and structure modification
- more accurate estimates of the effects of covariates
- more accurate estimates of the dose