UPPSALA
UNIVERSITET

UPTEC X 14 002
Examensarbete 30 hp
Februari 2014

Predictability and performance of different non-linear mixed-effects models for HbA1c in patients with type 2 diabetes mellitus

Gustaf Wellhagen

UPPSALA UNIVERSITET

Bioinformatics Engineering Programme

Uppsala University School of Engineering

UPTEC X 14 002	**Date of issue 2014-02**

Author

Gustaf Wellhagen

Title (English)

Predictability and performance of different non-linear mixed-effects models for HbA1c in patients with type 2 diabetes mellitus

Title (Swedish)

Abstract

To accurately predict the outcome of a late phase study, pharmacometric models can help in drug development. Making informed decision on which models to use will also facilitate drug development. This can depend on the mechanism of action for the drug as well as stability and runtime factors.

This is an investigation of four published semi-mechanistic pharmacometric models to predict glycosylated red blood cells (HbA1c) in a late phase study of an anti-diabetic drug together with an assessment of their stability and power to detect drug effects. Mean plasma glucose (MPG), fasting plasma glucose (FPG) or FPG and fasting serum insulin (FSI) are used together with HbA1c as drivers for change in the models. We find that less complex models, with fewer differential equations, are quicker to run and more stable, and that MPG alone is superior to FPG or FPG and FSI to detect a drug effect. The findings are useful for drug development in the anti-diabetic area, and show that a less mechanistic model performs well under these conditions.

Keywords

Type 2 diabetes mellitus, semi-mechanistic models, HbA1c, glucose, insulin, NONMEM

Supervisors

Dr. Maria Kjellsson
Uppsala University

Scientific reviewer

Dr. Andrew Hooker
Uppsala University

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **35**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 471 4687

Predictability and performance of different non-linear mixed-effects models for HbA1c in patients with type 2 diabetes mellitus

Gustaf Wellhagen

Populärvetenskaplig sammanfattning

Typ 2-diabetes är ett växande problem i världen och det beräknas att över 300 miljoner människor kommer att vara drabbade år 2030, vilket är dubbelt så många som år 2000. Det är främst i Asien och Afrika som det förutspås bli en folksjukdom. Förhöjda blodsockernivåer är det som karaktäriserar sjukdomen, vilka uppstår på grund av en kombination av minskad insulinproduktion, minskad insulinkänslighet och förhöjd leverglukosproduktion. De befintliga läkemedlen mot typ 2-diabetes inriktas på att kunna kontrollera blodsockernivåerna.

Farmakometrisk modellering som använder ickelinjära modeller med blandade effekter (fixerade eller slumpmässiga inom en population) används för att beskriva data från kliniska studier. Man vill kunna beskriva olika biomarkörer med hjälp av sådana modeller för att optimera kliniska studier inom läkemedelsutvecklingen. Genom detta kan man undvika att exponera patienter för toxiska doser, icke verksamma doser eller bestämma antalet patienter och mätningar som krävs för att uppnå signifikanta resultat och därmed spara pengar.

I den här studien jämförs fyra publicerade modeller för att beskriva HbA1c (andelen glykosylerade röda blodkroppar av det totala antalet röda blodkroppar). De är uppbyggda på olika sätt och använder olika typer av data för att göra sina prediktioner. Vi fann att modellerna som använder medelglukos är bättre än de som använder fasteglukos för att förutspå slutvärdet på HbA1c i simulerade studier. De är dessutom stabilare och bättre på att detektera läkemedelseffekter. Slutligen fann vi att modeller med färre differentialekvationer gav kortare körtid och bättre stabilitet.

Genom farmakometrisk modellering kan vi påskynda utvecklingen av nya läkemedel och förbättra den nedslående statistiken i de kostsamma kliniska faserna, där mer än hälften av alla läkemedelskandidater läggs ner.

Examensarbete 30 hp
Civilingenjörsprogrammet i bioinformatik
Uppsala universitet februari 2014

Index

Glossary of abbreviations

AUC	Area under curve
BID	Bis in die – twice daily
BIS	Basal insulin secretion
CLG	Clearance of glucose
CLGI	Insulin dependent clearance of glucose
EC50	Effective concentration to get 50% of maximum response
EGP	Endogenous glucose production
Emax	Maximal effect
FPG	Fasting plasma glucose
FSI	Fasting serum insulin
GKA	Glucokinase activator
HbA1c	Glycosylated haemoglobin
IGI model	Integrated glucose-insulin model
IGRH model	Integrated glucose-red blood cell-HbA1c model
IIV	Inter-individual variability
MCMP	Monte Carlo mapped power
MPG	Mean plasma glucose
MTT	Mean transit time
NONMEM®	Non-linear mixed-effects modelling software
OFV	Objective function value
OGTT	Oral glucose tolerance test
PPAR	Peroxisome proliferator-activated receptor
PsN	Pearl-speaks-NONMEM
QD	Quaque die – once daily
RBC	Red blood cells
T2DM	Type 2 diabetes mellitus

Introduction

Type 2 diabetes mellitus (T2DM) is a worldwide problem, today affecting around 150 million people (1). This number is expected to double by 2030. It is a metabolic disorder recognised through high blood glucose levels, which is the result of a combination of decreased insulin secretion, reduced insulin sensitivity and increased endogenous glucose production. T2DM is a lifelong disease and the standard of care is the glycaemic control drug metformin, combined with exercise and dietary advises (2).

Reducing the plasma glucose levels is the key to treating T2DM today (2). Thus, measuring these levels is important. Self-monitoring of blood glucose on a daily basis, is usually done by assessing fasting plasma glucose (FGP). However, these values vary greatly between occasions. The level of glycosylated haemoglobin, HbA1c, is a more long-term measurement of plasma glucose, which is commonly measured by the physician or the diabetes nurse.

Glucose has a natural tendency to bind to the haemoglobin of red blood cells (RBC) by a non-enzymatic process. This reaction is irreversible and the higher the glucose level in the blood is the larger will the fraction of the glycosylated RBC be. The ratio of the glycosylated RBC to the total RBC is called the HbA1c. As red blood cells have a life-span of about three months, the HbA1c will reflect the glucose levels in blood over the same duration. A healthy value of HbA1c is 4-6% and a level above 6.5% is a suggested basis for diagnosing diabetes (3).

The development of type 2 anti-diabetic drugs is made difficult because of the high variability in drug response between patients, the disease progression and other confounding effects. For instance, most people with T2DM are already on metformin or some other glycaemic control drug plus a diet and exercise schedule so isolating the effect of a new drug can be difficult. The pharmacometric modelling approach allows incorporating inter-patient, inter-occasion and population variability to overcome these problems.

Pharmacometric modelling uses non-linear mixed effect models to describe data from clinical trials. Such models incorporate a number of parameters, some of which are fixed effects, describing the main trend in the population, and some are random effects, describing the variability in the population or between observation occasions; hence *mixed effects*. Through these models one can make longitudinal predictions of biomarkers such as glucose, insulin or HbA1c.

By simulating clinical trials with varying dosing regimens one can optimise the study before actually conducting the trial. Optimised doses investigated in drug development will speed up the drug development programme, which saves money and minimises the risk of exposing patients to potentially toxic or non-efficacious drug concentrations. Optimising the right number of individuals to include in the study to get statistical significant results is another way of saving money and avoiding exposing too many patients to the drug. Altogether this will facilitate drug development in clinical stages (4).

The models investigated in this study are named after the main authors of the respective papers where they were first published. These are de Winter (5), Hamrén (6), Lledó (7) and Møller (8).

Aims

There are two main objectives in this study:

1. To investigate the predictability of four different published non-linear mixed-effects models for HbA1c in patients with T2DM.

2. To investigate the performance and stability of the four models to detect drug effects, looking at both existing and possible future targets for drug development.

By being able to describe these two objectives, both choosing the right model and deciding the study size for further clinical trials will be facilitated.

Methods

The main tools used in this project are different pharmacometric models executed through PsN (9,10) which runs NONMEM® 7.2 (11). Data handling and graphics were carried out in R (12). The models used in the project are further described below.

Models

The integrated glucose-insulin (IGI) model (13) is an established framework for describing the complex interplay between glucose and insulin in the body, developed for investigating glucose provocation studies. The version of the IGI model used in this project was for oral glucose tolerance test (OGTT) data, see Figure 1. There are some parameters describing the underlying system, some specific to the study and some specific to the drug. The model consists of compartments for glucose and insulin with transit compartments to describe absorption of glucose from food and effect compartments to describe delay of feedback mechanisms. A detailed description of the model is found in Jauslin *et al* (13).

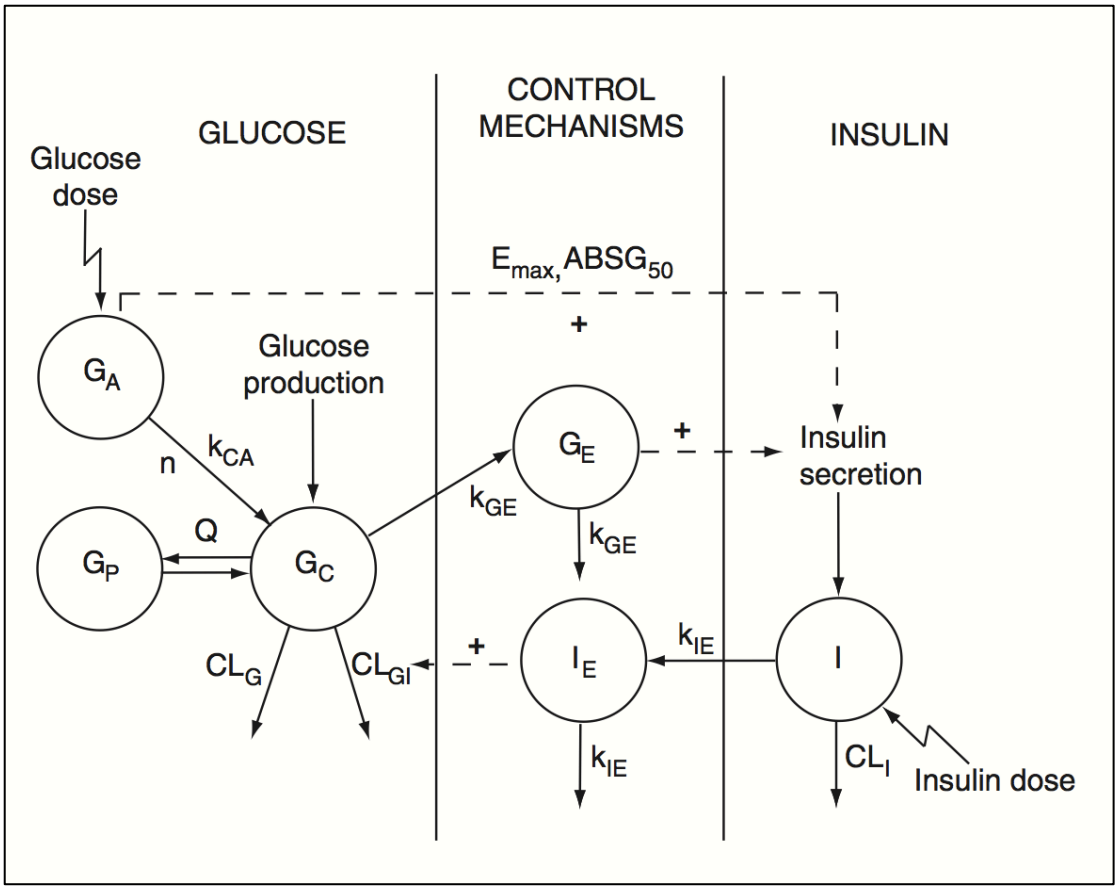

Figure 1. A schematic picture of the integrated glucose-insulin (IGI) model.

Schematic representation of the oral glucose tolerance test (OGTT) model. Full arrows indicate flows, and broken arrows indicate control mechanisms. G and G , central and peripheral compartments of glucose; G_A, representation of the transit compartments for glucose absorption; G_E, effect compartment of glucose for the control of insulin secretion; I, insulin disposition compartment; I_E, effect compartment of insulin for the control of glucose elimination; Q, CL_G, CL_{GI}, k_{CA}, n, kinetic parameters of the glucose submodel; CL_I, insulin clearance; k_{GE} and k_{IE}, rate constants for the effect compartments; E_{max}, maximal effect of the glucose absorption rate on insulin secretion; $ABSG_{50}$, glucose absorption rate producing 50% of E_{max}. Used with permission.[i]

The concentrations of insulin and glucose are influenced by each other through feedback mechanisms but also by drug effects when drug concentrations are present. The IGI model allows for six different drug effects (some of which are hypothetical) to be implemented. These are:

- **Absorption,** where the absorption of glucose into the blood is decreased due to a drug inhibiting the breakdown of glucose polymers, for example alpha-glucosidase inhibitors.

- **Basal insulin secretion (BIS),** where the endogenous production of insulin is increased, such as sulfonylureas.

- **Clearance of glucose (CLG),** where the renal clearance of glucose from the blood is increased, such as SGLT2 inhibitors.

- **Endogenous glucose production (EGP),** where the hepatic production of glucose is decreased, such as biguanides.

- **Incretin effect,** where the release of insulin is stimulated after elevated blood glucose levels, such as GLP-1 analogs.

- **Insulin dependent clearance of glucose (CLGI),** where the renal clearance of glucose from the blood is increased in relation to the concentrations of insulin. Thiazolidinediones has been hypothesized to have this effect, however this class of drug has several mechanisms of action (14).

The de Winter model (5) consists of linked indirect response models describing HbA1c, FPG and fasting serum insulin (FSI), see Figure 2. This model is developed to describe the disease progression of T2DM with respect to beta cell function and insulin sensitivity, in a semi-mechanistic manner. There is a homeostatic feedback between FSI and FPG, and FPG also affects the HbA1c. Also this model includes system-specific parameters as the insulin sensitivity (S) and the beta cell function (B). Also in this model can drug effects be included on various sites depending on the mechanism of action of the drugs. Drugs with action on BIS and incretin would affect B, glucose absorption and EGP would affect the input to FPG, CLG would affect output from FPG while CLGI would affect S. A detailed description of the model is available in de Winter *et al* (5).

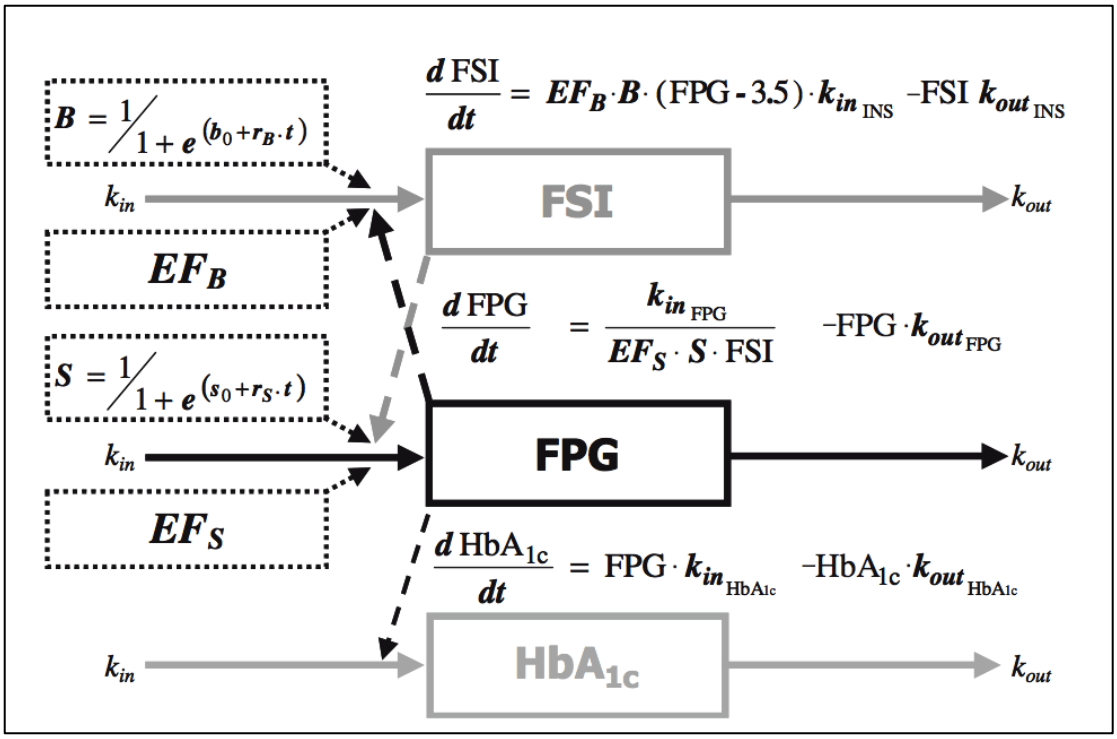

Figure 2. A schematic picture of the de Winter model.

Schematic representation of the structure of the mechanism-based population PD disease progression model, including the homeostatic feedback between FSI and FPG and the feed-forward between FPG and HbA_{1c}. Used with permission.[ii]

The Hamrén model (6) focuses at RBC in 8 compartments; 4 for non-glycosylated and 4 for glycosylated RBC, see Figure 3. In their life-span they can by glycosylated to HbA1c through FPG~. The FPG is modelled using an indirect response model and the FPG~ is created by raising FPG to a power, allowing the HbA1c production to be non-linear with FPG. This model also has a sex difference incorporated on the RBC life-span. The distribution in the different dose arms in the clinical study was used when randomly assigning sex to subjects. In this model the drug effect was incorporated on the output of FPG. A detailed description of the model is available in the paper by Hamrén *et al* (6).

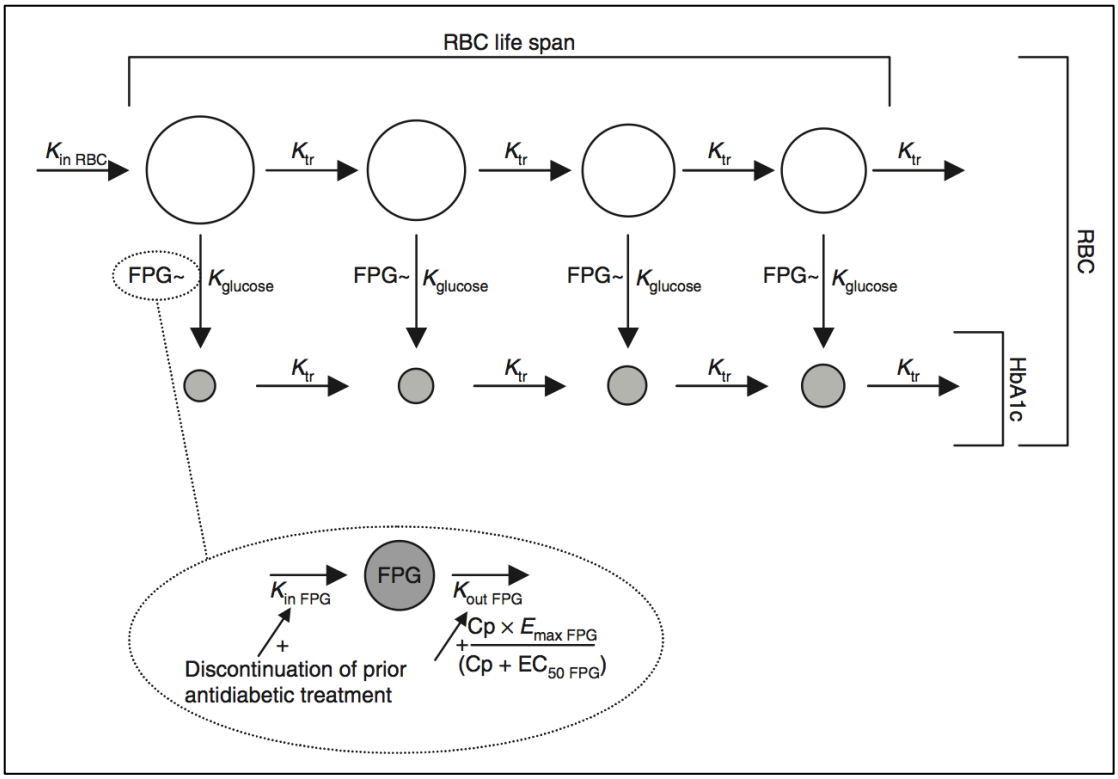

Figure 3. A schematic picture of the Hamrén model.

Schematic representation of the mechanism-based model for the FPG–HbA1c relationship. Cp, tesaglitazar plasma concentration; $EC_{50\ FPG}$, tezaglitazar plasma concentration achieving half-maximal effect on $E_{max\ FPG}$; $E_{max\ FPG}$, maximum effect on $K_{out\ FPG}$, FPG, fasting plasma glucose; Hb, hemoglobin; HbA1c, glycosylated hemoglobin; $K_{glucose}$, glycosylation rate constant of RBCs to HbA1c; $K_{in\ RBC}$, zero-order release constant of RBCs into the circulation; $K_{in\ FPG}$, zero-order rate constant for the production of FPG; $K_{out\ FPG}$, first-order rate constant for the removal of FPG from the blood; K_{tr}, first-order transit rate constant; RBCs, red blood cells. Used with permission.[iii]

The Lledó model (7) is similar to the Hamrén model in its structure and focusing on RBC and their glycosylation over their life-span, see Figure 4. There are 12 compartments for glycosylated and 12 for non-glycosylated RBC. In this model the mean plasma glucose (MPG) governs the glycosylation of RBC in a linear fashion. To reduce runtimes in the second part of this project the model was reduced to 6+6 compartments for RBC. This model includes a glucose effect on the RBC life-span; the higher the glucose the shorter the life-span. The MPG was modelled using an indirect response model and drug effects were included on the output of MPG.

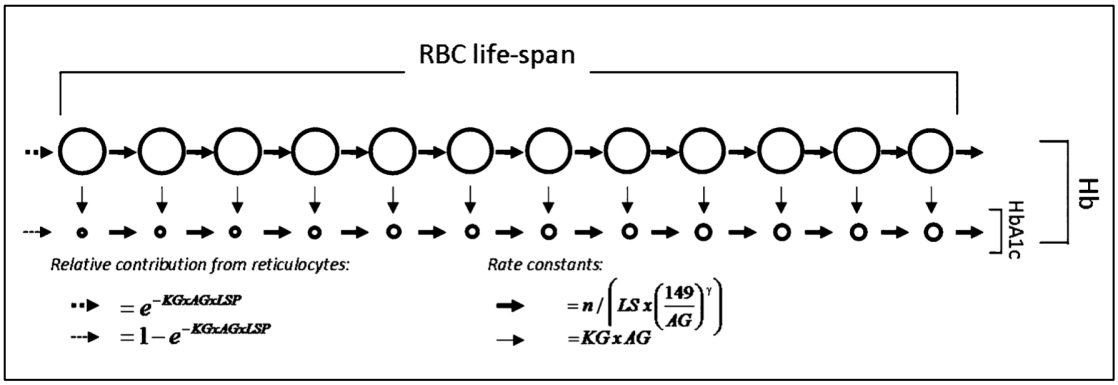

Figure 4. A schematic picture of the Lledó model.

Illustration of the final IGRH model. The *areas of the circles* represent the relative amount of RBC at the different stages under constant RBC production. *n* represents the number of transit compartments which is fixed to 12. Used with permission.[iv]

The Møller model (8) also consists of indirect response models (see Figure 5) one for MPG and one for HbA1c, where the production of HbA1c is driven by MPG affecting the input of the HbA1c model. Drug effects are driving the MPG to a post-treatment MPG level by affecting the input to the MPG model. A detailed description of the model is available in the paper by Møller *et al* (8).

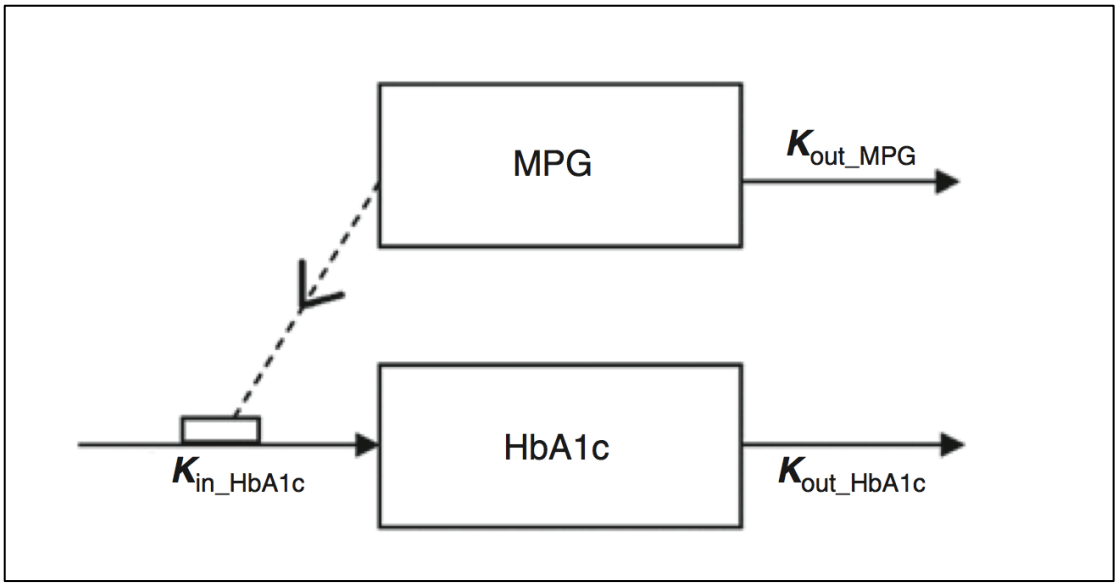

Figure 5. A schematic picture of the Møller model.

The model is an indirect response model where the production of HbA1c is stimulated by mean plasma glucose (MPG) through the parameter k_{in_HbA1c} that is fixed to 0.081%/mmol/l per week. The model is initialized in steady state, at the time of the screening visit, where MPG_{ss} is the value for MPG. MPG is assumed to change during a washout/run-in period toward MPG_{base} typically obtained at the baseline visit. $MPG_{posttreatment}$ is the stable glucose value obtained after introducing the experimental treatment. k_{out_MPG} is the rate constant defining the rate of treatment onset on MPG. The parameter k_{out_HbA1c} defines the output rate constant for HbA1c and is fixed to 0.226 per week. The present model further introduces a parameter β that allows an offset in the linear relationship between MPG and HbA1c in steady state. Thus, k_{in_HbA1c} is stimulated by MPG + β. Used with permission.[v]

Approach 1 – Predictability through simulation studies

For investigating predictability, a phase II study was simulated with 210 individuals evenly distributed in 6 dose arms; placebo, 25 mg twice daily (BID), 50 mg once daily (QD), 50 mg BID, 100 mg QD and 100 mg BID of a glucokinase activator (GKA). The drug effects and parameter values were taken from the publication Kjellsson *et al* (4). Daily glucose intake was assumed to be 3 large meals and 3 snacks in between the large meals. The GKA was administered in the morning 30 minutes prior to breakfast (QD and BID) and 30 minutes prior to dinner (BID). Initial glucose values were set to match the inclusion criteria of the study. The study was 12 weeks (84 days) long. To allow the glucose levels to stabilize before drug administration a run-in period of 1 week was used. The run-in period was omitted from the study for the results.

The IGI model was run to produce glucose and insulin data. The GKA drug has two effects that were implemented: it lowers the EGP and increases the BIS. Five hundred replicates of the study were performed.

The FPG, FSI and MPG were extracted from the data simulated by the IGI model and used in the de Winter model (FPG, FSI), the Hamrén model (FPG), the Lledó model (MPG) and the Møller model (MPG) to simulate HbA1c. The FPG and FSI

were assessed as the glucose and insulin concentrations before the first meal in the morning while MPG was calculated as the mean of each day, meaning AUC/24. The placebo and baseline corrected HbA1c was then compared with the outcome of the study that was performed by Hoffman-La Roche (4).

The aim for this part was to investigate the final predictions of HbA1c and the overall fit to the experimental data across the four models.

Approach 2 – Power calculations

To investigate the power to detect a drug effect plus stability and runtimes of the models a hypothetical study was simulated. Four dose arms were used, with individuals evenly distributed in: Placebo, 25 mg BID, 50 mg BID and 100 mg BID.

The IGI model was modified to incorporate the six drug effects separately and used for simulating FPG, FSI and MPG. As the Lledó model is the most mechanistic of the four models regarding formation of HbA1c, it was used to simulate HbA1c observations using the MPG from the IGI runs.

The aim was to investigate which model is the best at detecting the drug effect and assess how many individuals would be needed to maintain power in a clinical study. The hypothetical drug effect was titrated to give a difference in 10% in AUC compared to placebo for the highest dose arm; values for the corresponding Emax and EC50 are shown in Table 1. The drug effect was assumed to be proportional and was calculated as 1 - Emax*C/(EC50+C) for drug effects on absorption and EGP or 1 + Emax*C/(EC50+C) for drug effects on BIS, CLG, CLGI and incretin.

Table 1. Emax and EC50 values used to titrate the drug effect to a 10% drop in HbA1c for the highest dose.

Drug effect	Emax	EC50
Glucose absorption	0.75	0.03
BIS	3.5	0.5
CLG	3	0.3
EGP	1	0.2
Incretin	2.5	0.1
CLGI	2	0.1

The Monte Carlo Mapped Power method (MCMP) is a quick algorithm for finding the number of patients in a clinical study required to get a statistically significant power to detect a drug effect (15). In short, the method needs two model files: a full and a reduced model. Instead of performing repeated estimations for data set with increasing number of individuals as traditionally is done when assessing power, two estimations with a very large data set are performed, individuals are sampled from the data set with replacement and increasing sample size until full power is achieved or all individuals are included and the difference in objective function value (OFV) between the full and reduced model is calculated. The power at different significance levels is assessed through a chi-squared

distribution decided by the degrees of freedom. Since the drug effect was modelled using Emax-models with two parameters, EC50 and Emax, the degrees of freedom were set to 2 in this study. In Table 2 are listed the fixed parameters and their respective values used in the MCMP runs.

Table 2. A list of the fixed parameters and their values used for the MCMP runs.

Model	Parameter	Value
de Winter full		
	K_{OUT} FPG	1
	Emax	
de Winter reduced		
	K_{OUT} FPG	1
	Emax	0
	EC50	53
Hamrén reduced		
	Emax	0
	EC50	1
Lledó full		
	K_{IN} HbA1c	1
	K_{OUT} MPG	0.226
	Exp Glucose-Hb	0.381
	Life-span RBC	91.7
	IIV LS	0.0822
	LSP	8.20
	IIV LSP	0.115
Lledó reduced		
	K_{IN} HbA1c	1
	K_{OUT} MPG	0.226
	Exp Gluc-Hb	0.381
	Life-span RBC	91.7
	IIV LS	0.0822
	LSP	8.20
	IIV LSP	0.115
	Emax	0
	EC50	1
Møller full		
	K_{OUT}	0.226
	K_{IN}	0.081
Møller reduced		
	K_{OUT}	0.226
	K_{IN}	0.081
	Posterior glucose	0
	EC50	35

To reduce runtime and increase stability, the Lledó model was reduced to 13 differential equations for this part; 1 for MPG, 6 for non-glycosylated RBC and 6 for glycosylated RBC.

As the change of FPG and FSI is fast compared to HbA1c, the de Winter model could be reduced to a steady state solution for FPG and FSI with one differential equation of the HbA1c. The number of parameters estimated for each model in the MCMP runs is shown in Table 3, where the two extra parameters in the full model are Emax and EC50.

Table 3. Number of parameters estimated for each model.

Model	Full model	Reduced model
de Winter	15	13
Hamrén	13	11
Lledó	7	5
Møller	11	9

The standard simulation setup was defined as sampling on weeks 0, 4, 6, 8 and 12. To investigate the effect of sparser sampling, also the schemes of sampling on weeks 0, 4 and 8 or weeks 0, 6 and 12 were tested, where sampling up to 8 weeks also represents performing the study for a shorter duration. To investigate the power to detect a change from placebo with a lower power, one setup was defined using the placebo and the lowest dose arm. The simulation setup is summarised in Table 4.

Table 4. Summary of the simulation setups with different arms, sampling schedules and study durations.

Simulation set up	Arms	Sampling schedule	Duration of study
Standard	0, 25, 50, 100	0, 4, 6, 8, 12	12
Fewer samples	0, 25, 50, 100	0, 6, 12	12
Shorter duration	0, 25, 50, 100	0, 4, 8	8
Fewer dose arms	0, 25	0, 4, 6, 8, 12	12

Initial estimates in the models were continuously updated throughout the investigation to increase the probability of a successful run.

Results

Predictability

The 95% confidence intervals around the mean for the baseline and glucose corrected HbA1c for the different models stratified on dose arm are shown in Figure 6, Figure 7, Figure 8 and Figure 9. The true values from the clinical study by Hoffman-La Roche are included.

The de Winter model predicted a greater decrease in HbA1c than the clinical data shows. The slopes for all dose arms seemed to level out by the end of 12 weeks.

The Hamrén model predicted a slower time to reach the new HbA1c than the other models and that dosing twice daily seemed to have a greater effect than once daily dosing schedule.

The Lledó model overpredicted the drop in HbA1c in all scenarios.

The Møller model overpredicted greatly for the 50 mg BID dose arm, but for the other dose arms it was less off.

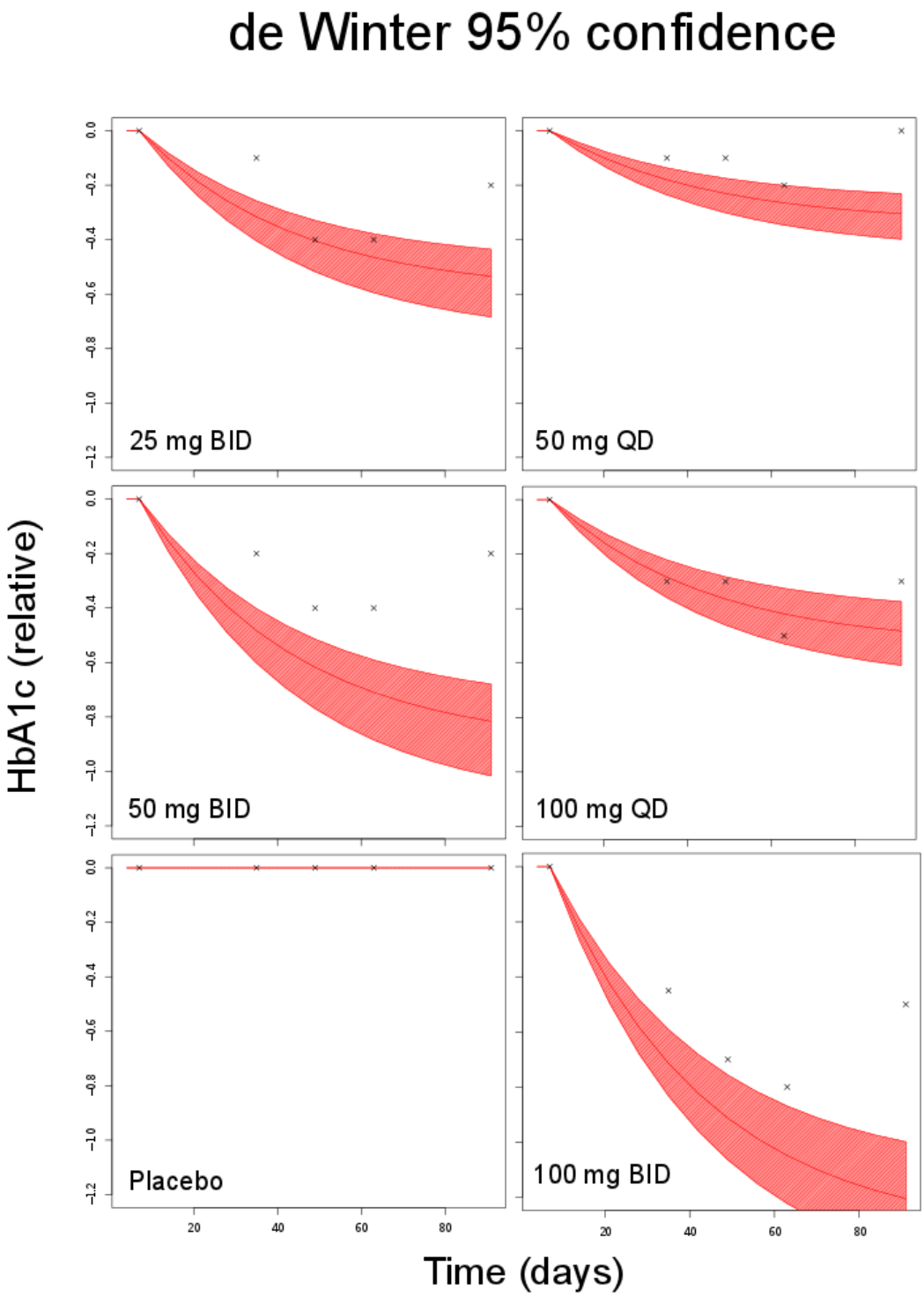

Figure 6. Longitudinal predictions of baseline and placebo corrected HbA1c. The 95% confidence interval around the predicted mean of the simulated data for the de Winter model is shown as the shaded area. The crosses are the clinical data.

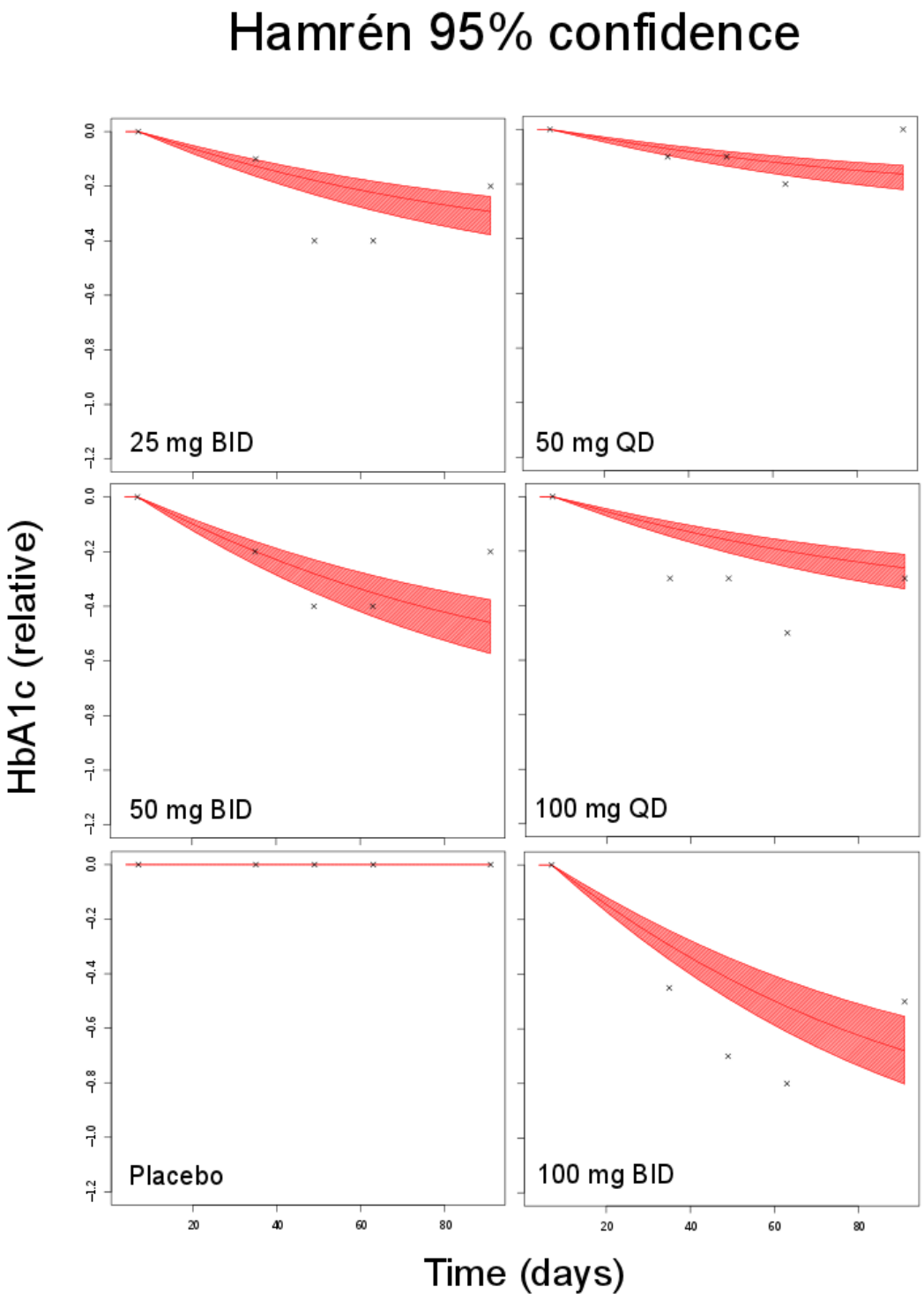

Figure 7. Longitudinal predictions of baseline and placebo corrected HbA1c. The 95% confidence interval around the predicted mean of the simulated data for the Hamrén model is shown as the shaded area. The crosses are the clinical data.

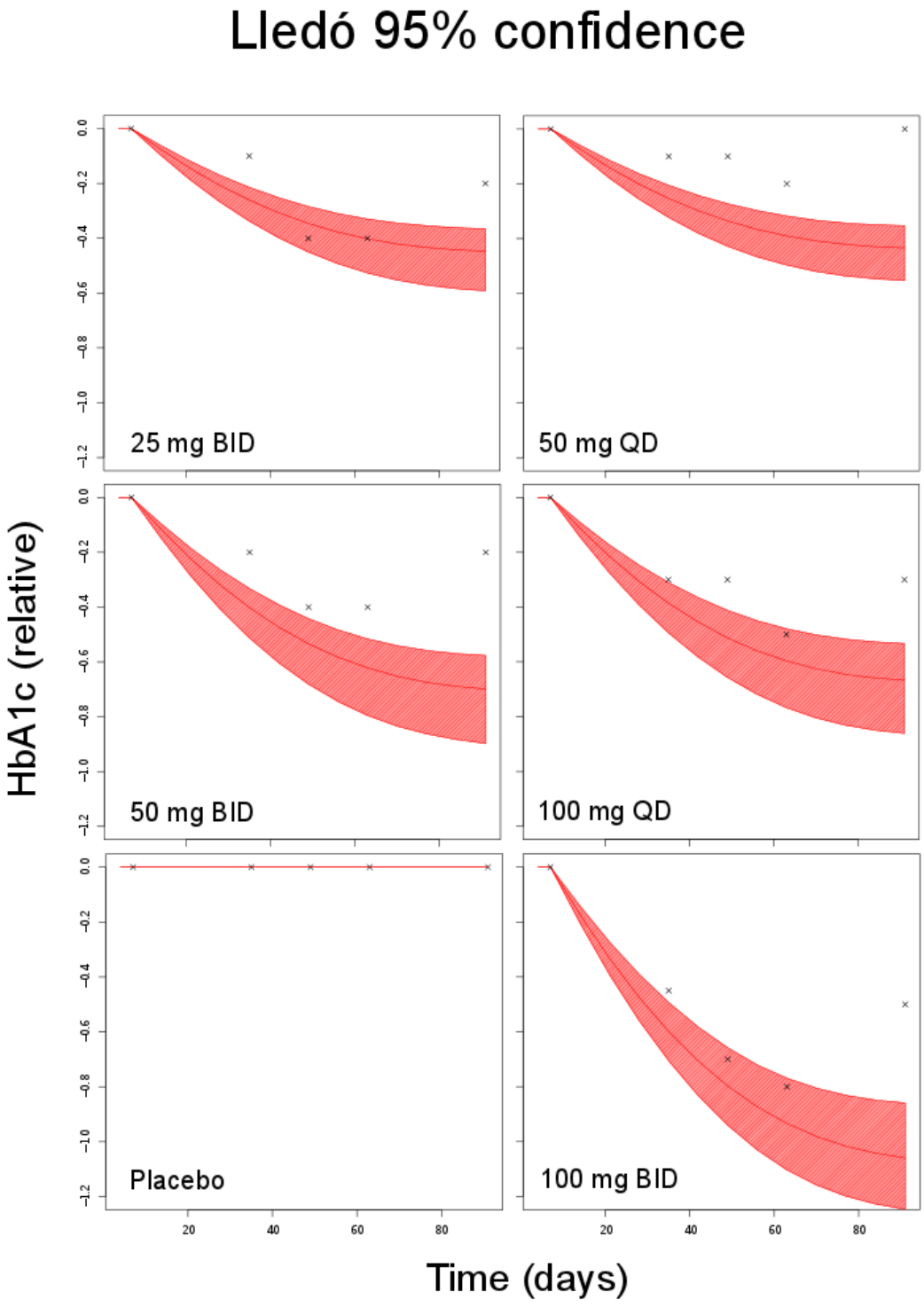

Figure 8. Longitudinal predictions of baseline and placebo corrected HbA1c. The 95% confidence interval around the predicted mean of the simulated data for the Lledó model is shown as the shaded area. The crosses are the clinical data.

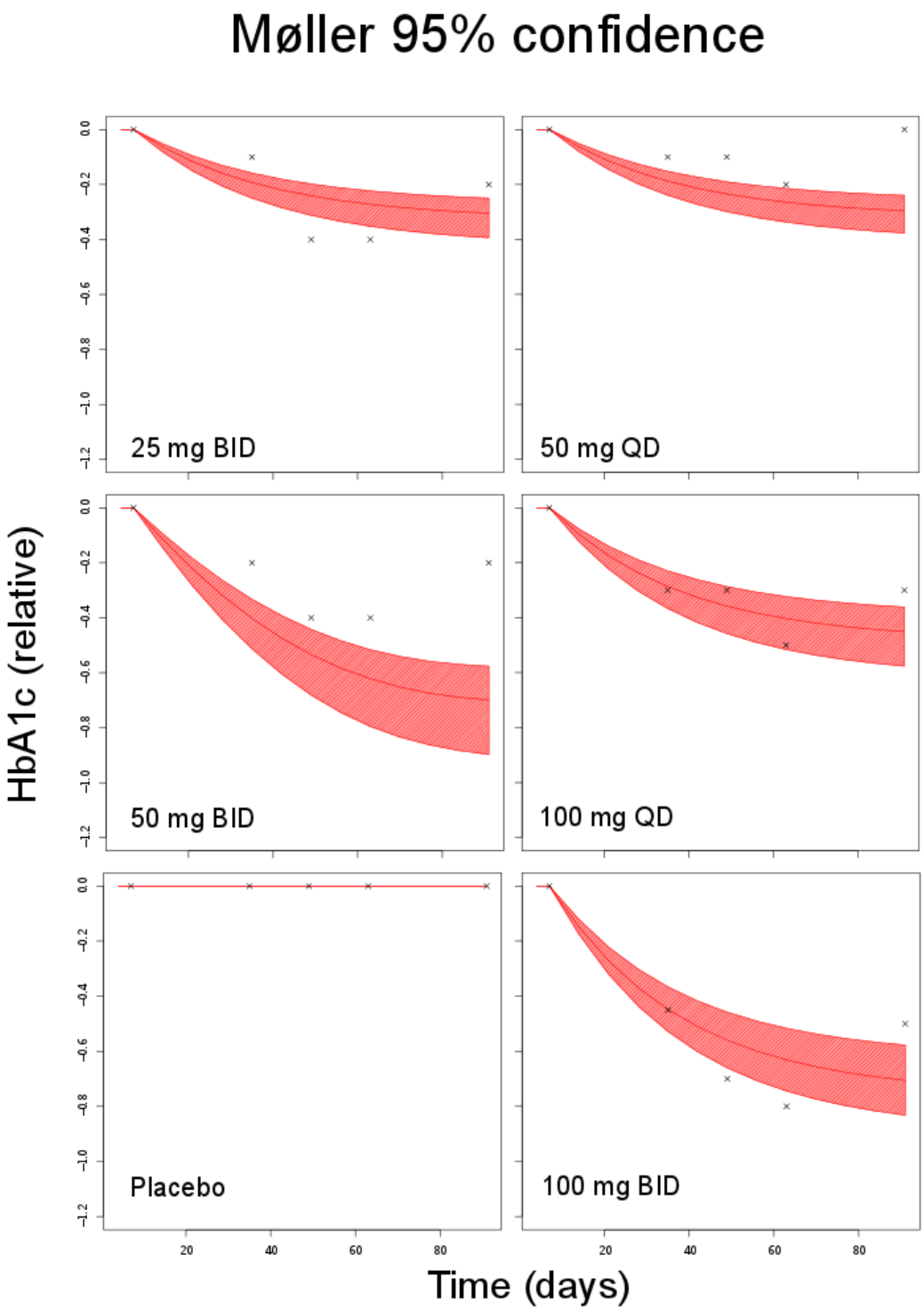

Figure 9. Longitudinal predictions of baseline and placebo corrected HbA1c. The 95% confidence interval around the predicted mean of the simulated data for the Møller model is shown as the shaded area. The crosses are the clinical data.

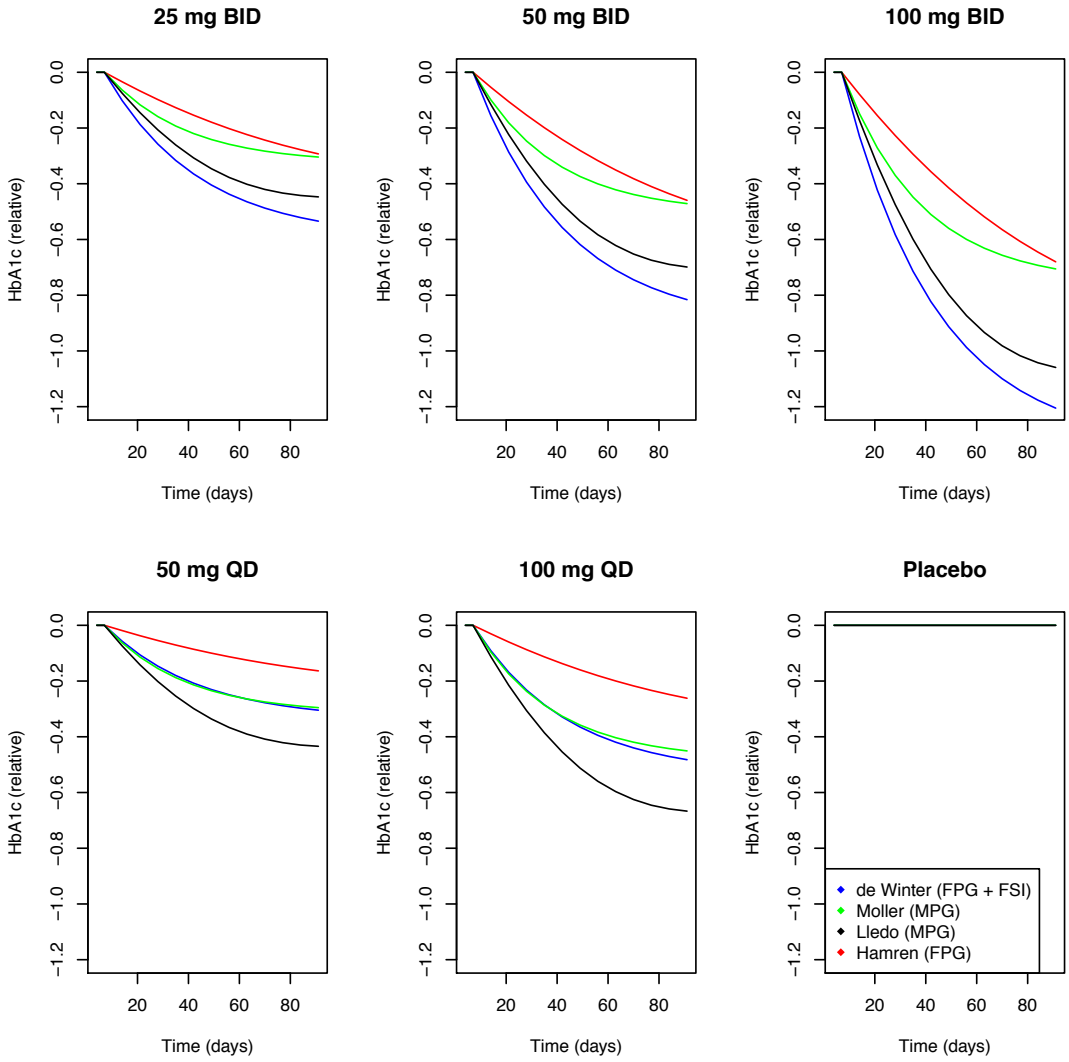

Figure 10. Longitudinal predictions of baseline and placebo corrected HbA1c by dose arm for the four HbA1c models. The first row shows plots for dosing BID and the second row for dosing QD and placebo.

In Figure 10, the mean prediction of the relative change in HbA1c for all the models for each dose arm is plotted. The first row with dosing BID and increasing doses shows a similar pattern, where the de Winter model predicts the greatest drop, while the second row with dosing QD indicates that the Lledó model predicts the greatest drop. Also, the HbA1c + MPG driven models (Lledó and Møller) have similar predictions in the QD arms, but diverged in the BID arms.

Stability

Stability was assessed using the termination message from NONMEM® with minimisation terminated used as a definition of crash. When taking into account the four different setups and six drug effects investigated (totalling 24 combinations), it is evident that the Møller and Hamrn models are superior in stability, see Table 5.

Table 5. Stability of all setups and all drug effects assessed through number of runs with results and successful runs. Success rates are given in parentheses.

Model	Runs with results	Successful runs
de Winter	16 (67%)	16 (67%)
Hamrén	23 (96%)	23 (96%)
Lledó	5 (21%)	2 (8%)
Møller	24 (100%)	23 (96%)

It should be noted that with most of the runs with the Lledó model, the full run ends up at an OFV value significantly better than the reduced even though the power calculation cannot be performed. For all models, the problems occurred more frequently with the full model file run than with the reduced model file.

Runtimes

In Table 6 and Table 7 the runtimes for the standard setup (with sampling on week 0, 4, 6, 8 and 12) is shown for two of the drug effects. As the runtimes are depending on how close the initial values are to the final estimates, it was difficult to find any clear trends. As expected, the run time with the full models was in most cases longer than the runtimes for the reduced model, as this model contains more parameters to estimate, see Table 3. Overall, the runtimes for the Hamrén model were long while the Møller model was quite quick to run.

Table 6. Runtimes in (hours:minutes:seconds) for standard setup runs of the drug effect on basal insulin secretion.

Model	Full	Reduced
de Winter	0:50:03*	0:47:44*
Hamrén	11:18:14	5:47:28
Lledó	17:55:12	15:05:32
Møller	3:38:05	0:37:14

*crash

Table 7. Runtimes in (hours:minutes:seconds) for standard setup runs of the drug effect on endogenous glucose production.

Model	Full	Reduced
de Winter	3:23:37	1:43:14
Hamrén	4:43:17	3:11:50
Lledó	0:01:46*	0:15:11
Møller	0:54:01	0:20:39

*crash

Performance

As seen in Figure 11, in the standard settings, the Møller and Lledó models have superior power to detect the drug effect over the de Winter and the Hamrén model. This is most likely related to Møller and Lledó models being driven by MPG while de Winter and Hamrén are driven by FPG. The Lledó model crashed however for two of these runs, and is thus not presented in Figure 11.
Overall, the models had the highest power to detect a drug effect on CLGI with the study design in the standard setting, followed by a drug effect on EGP.

The Hamrén model was the least powerful to detect a drug effect on glucose absorption or the incretin effect, which is related to this model using FPG as the driver of HbA1c formation. The de Winter model, which also uses FPG, performed slightly better in these drug effects, indicating that usage of FSI does contribute with important information for these drug effects. For the remaining drug effects: BIS, CLG, EGP and CLGI, the de Winter model was the least powerful, indicating that for these drug effects the FSI contributes only to a small extent and the mechanistic HbA1c formation expressed in the Hamrén model is of more importance..

Figure 12 shows the results from the setting with only one drug arm. As expected, the power of the models to detect the drug effect was lower for this design and more runs crashed and could not be displayed in the figure.

Figure 13 shows the results from the setting with study duration of only 8 weeks. Again, as expected the power of the models to detect the drug effect was lower than with study duration of 12 weeks.

In Figure 14 the results of the setting with fewer sampling points (three instead of five) are shown. The power to detect the drug effect is hardly affected by the reduction of sampling points.

A summary of the number of individuals needed to get 95% power in the study for each model and drug effect is shown in Table 8.

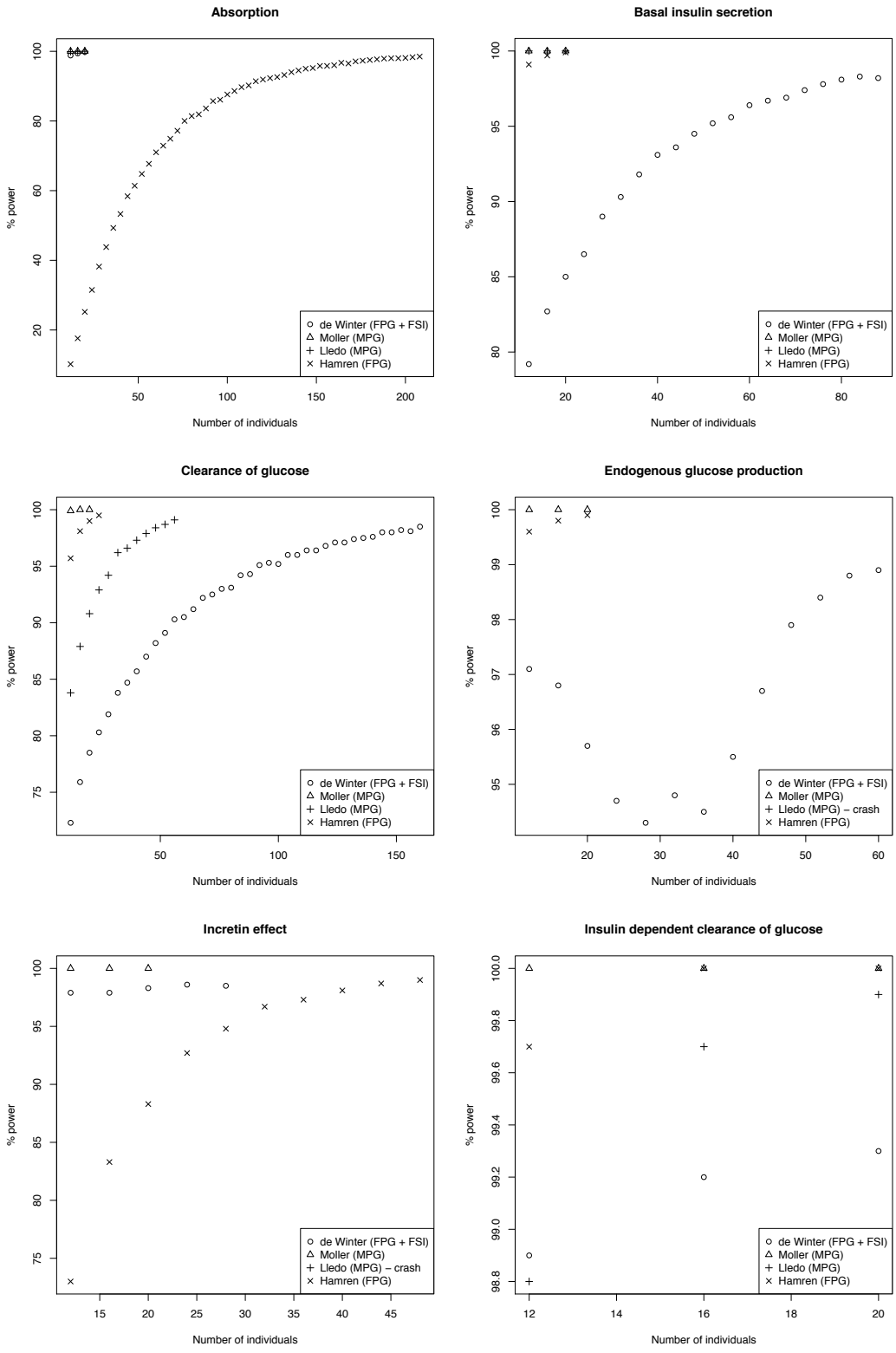

Figure 11. Power for the standard setup (sampling weeks 0, 4, 6, 8 and 12) against number of individuals. The drug effects are a) absorption, b) basal insulin secretion, c) insulin independent clearance of glucose, d) endogenous glucose production, e) the incretin effect and f) insulin dependent clearance of glucose. Note that the axes are varying between the panels.

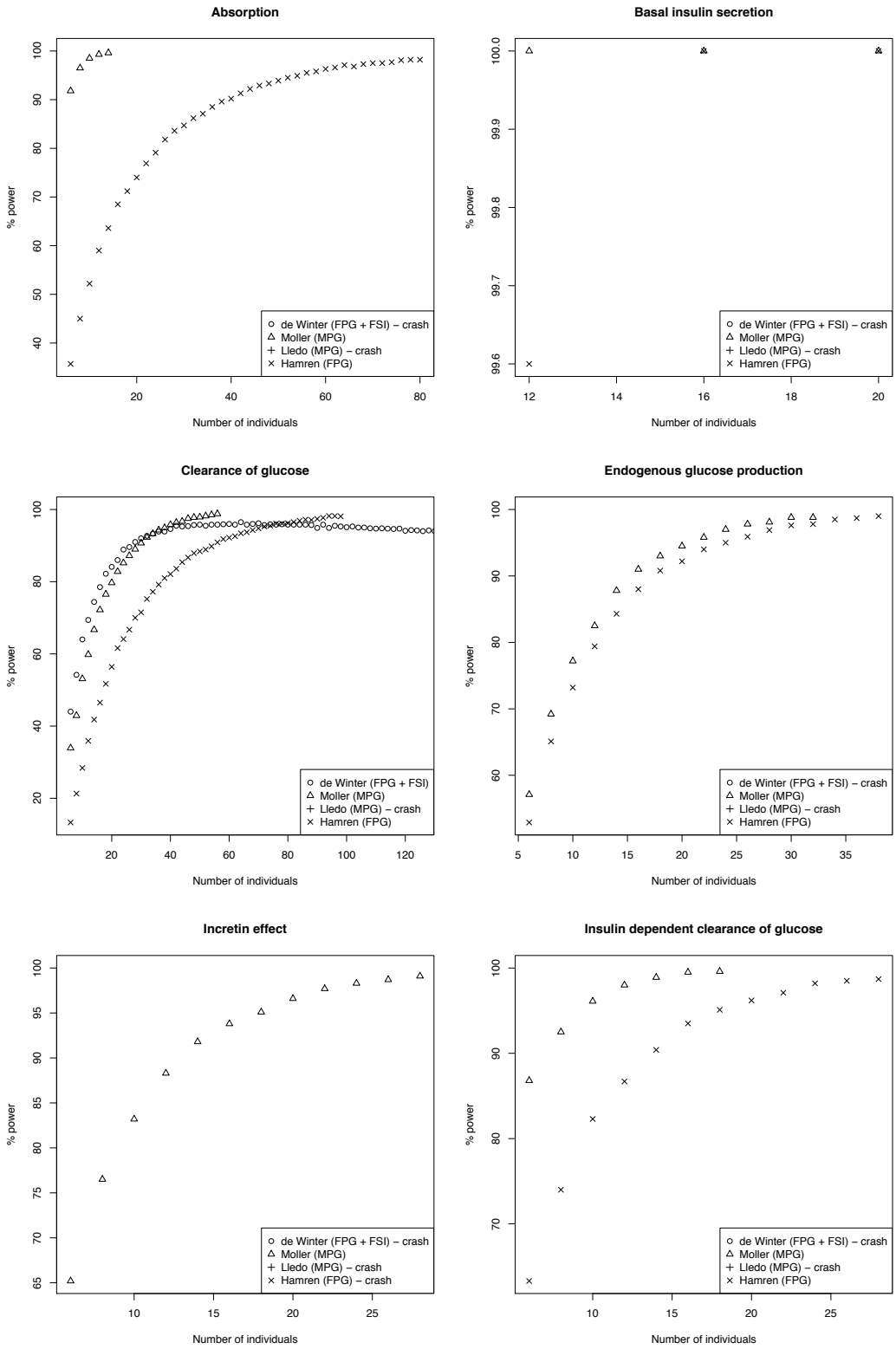

Figure 12. Power for the fewer dose arms setup (placebo + dose arm 25 mg BID, sampling weeks 0, 4, 6, 8 and 12) against number of individuals. The drug effects are a) absorption, b) basal insulin secretion, c) insulin independent clearance of glucose, d) endogenous glucose production, e) the incretin effect and f) insulin dependent clearance of glucose. Note that the axes are varying between the panels.

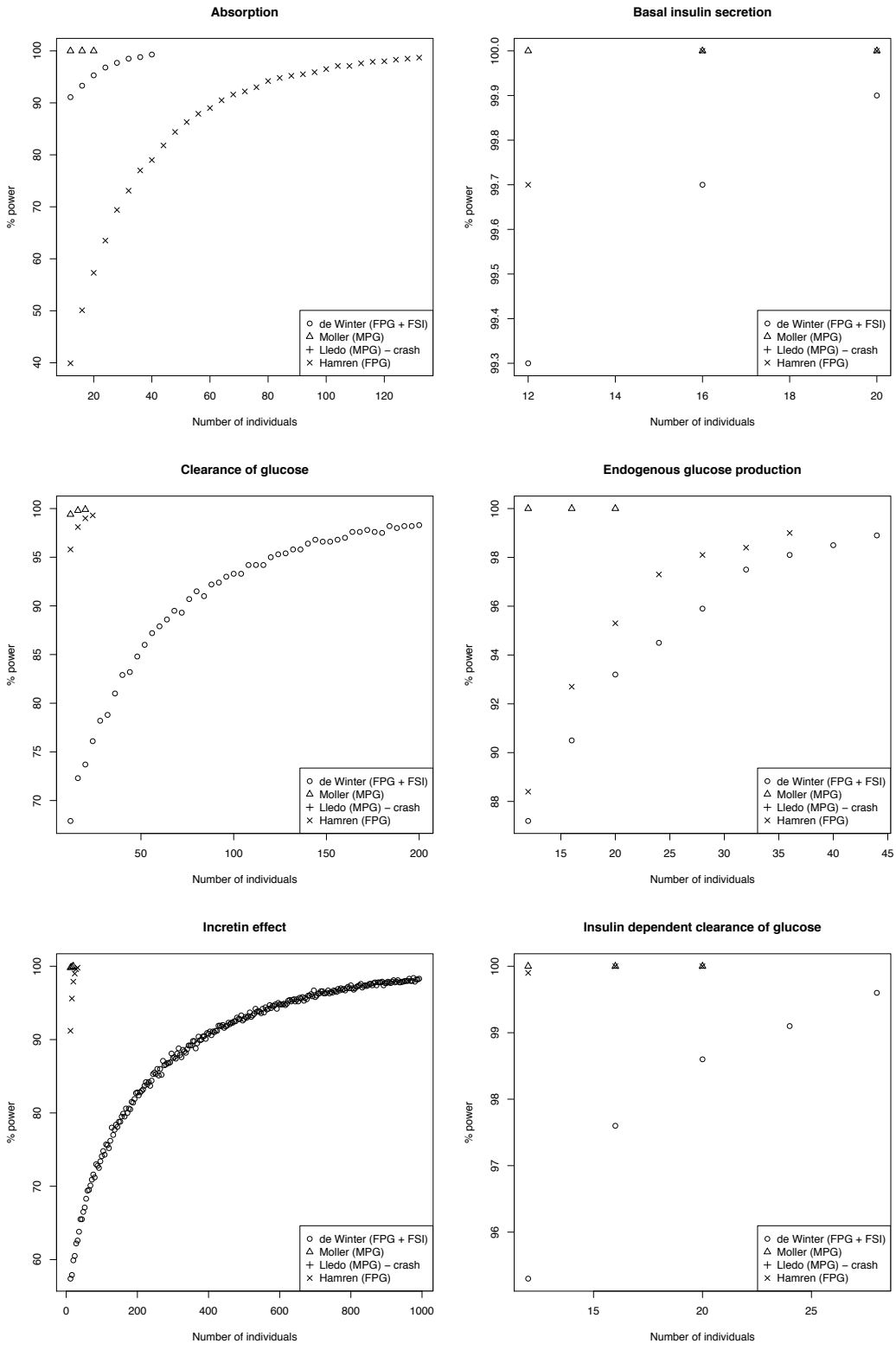

Figure 13. Power for the shorter study setup (sampling weeks 0, 4 and 8) against number of individuals. The drug effects are a) absorption, b) basal insulin secretion, c) insulin independent clearance of glucose, d) endogenous glucose production, e) the incretin effect and f) insulin dependent clearance of glucose. Note that the axes are varying between the panels.

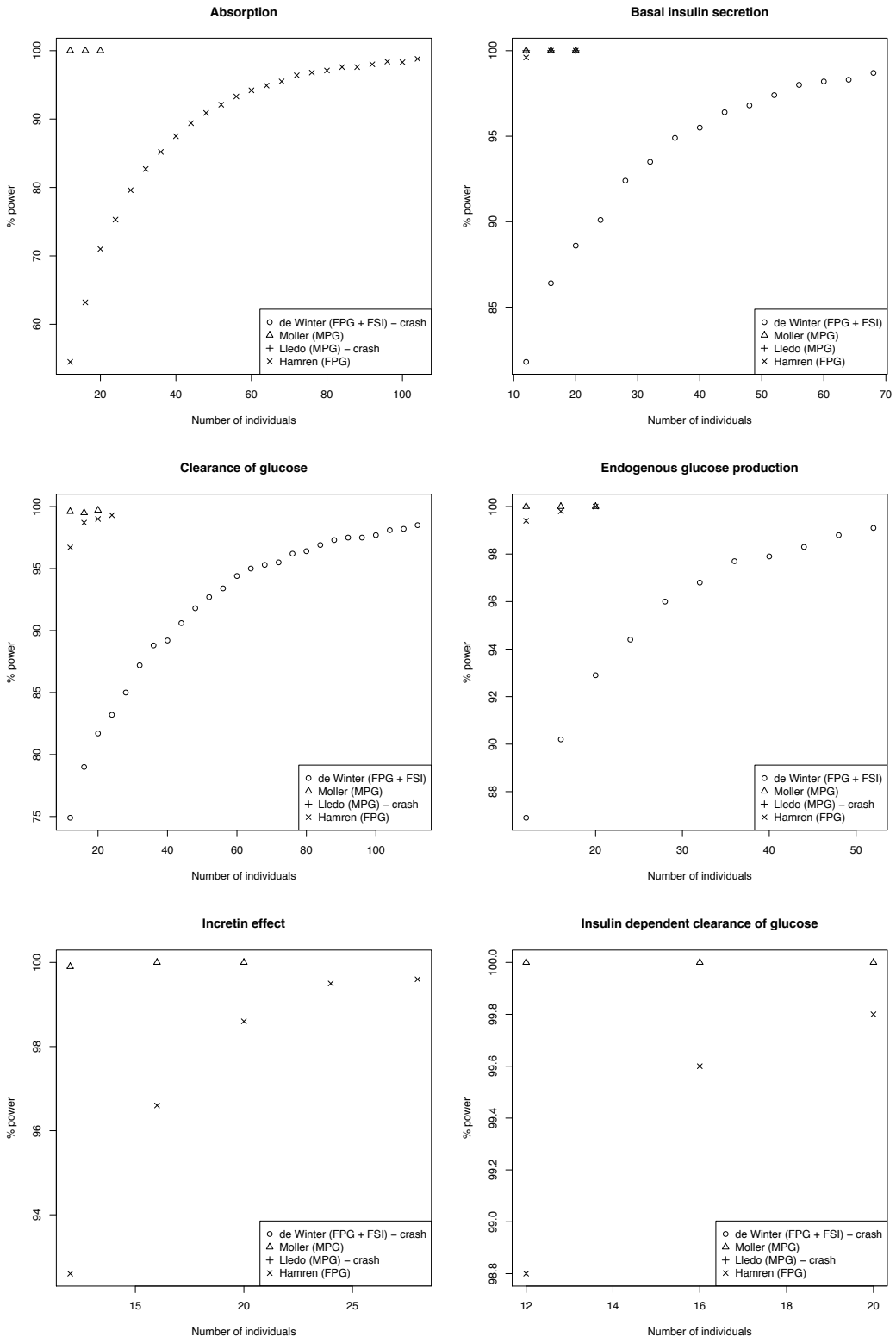

Figure 14. Power for the fewer samples setup (sampling weeks 0, 6 and 12) against number of individuals. The drug effects are a) absorption, b) basal insulin secretion, c) insulin independent clearance of glucose, d) endogenous glucose production, e) the incretin effect and f) insulin dependent clearance of glucose. Note that the axes are varying between the panels.

Table 8. The number of individuals needed to get 95% power in a study for each model and drug effect.

Setup	de Winter	Hamrén	Lledó	Møller
Absorption				
Standard	4	140	4	4
Fewer dose arms	*	52	*	8
Shorter duration	20	100	*	4
Fewer samples	*	60	*	4
BIS				
Standard	56	4	4	4
Fewer dose arms	*	2	*	2
Shorter duration	4	4	*	4
Fewer samples	36	4	4	4
CLG				
Standard	100	4	32	4
Fewer dose arms	40	80	*	40
Shorter duration	120	4	*	4
Fewer samples	68	4	*	4
EGP				
Standard	40	4	*	4
Fewer dose arms	*	2	*	2
Shorter duration	28	20	*	4
Fewer samples	28	4	*	4
Incretin				
Standard	4	28	*	4
Fewer dose arms	*	*	*	18
Shorter duration	624	16	*	4
Fewer samples	*	16	*	4
CLGI				
Standard	4	4	4	4
Fewer dose arms	*	20	*	10
Shorter duration	4	4	*	4
Fewer samples	*	4	*	4

* crash

Discussion

Predictability

The predictions from the various HbA1c models differed a little. All models, but the Hamrén model reached a new steady-state of the HbA1c before the end of the study. As no parameters were re-estimated from published values the time to reach steady-state for the models would be highly dependent on the study design and the drug tested in the original publication. The Hamrén model was developed using tezaglitazar, a peroxisome proliferated-activated receptor (PPAR) agonist, which alters whole body metabolism. An explanation for the long time to reach steady-stated could be that even though glucose has reached steady-state the HbA1c formation is altered by other mechanisms and would thus be slower than the RBC life-span.

No placebo model was used, even though it is known that patients will improve through partaking in a study of this kind. Diet and exercise, plus stricter compliance to the drug regimen (i.e. metformin for the placebo patients) are the main factors driving this change. Implementing a placebo model would make the simulations more realistic and overall reduce the power to detect a drug effect. To make the simulations even more realistic, a less strict meal schedule could be implemented. The simulations were now performed assuming 3 large meals and 3 snacks for all people of the exact same size with no deviation of meal time.

The last point (at week 12) in the clinical data shows deterioration from the previous measurement (week 8). This could be a wearing-off of the drug effect, which was not incorporated in our models but has been reported for previous substances in the GKA family (16) and might be a reason to discontinue the drug development. Another improvement might be to investigate if the two drug effects of the GKA are additive, which was modelled but might not be the case.

If the last data point is disregarded, the models would generally be closer to predicting the reported final HbA1c from the clinical study. In all, the Hamrén model shows the worst fit as it assumes an almost straight line in the decrease of HbA1c.

Stability

The Lledó model is very sensitive to initial estimates and only 4 out of 48 MCMP runs could be finished with satisfactory results. Since HbA1c observations were created using the Lledó model, one would think that the MCMP would be biased towards that model. Indeed, if only looking at the drop in OFV, it seems to be performing very well but nevertheless it has troubles finishing the runs, something that can perhaps be ascribed to the complex structure of the model.

The de Winter model also seems to be sensitive to initial estimates, as it does not complete with a successful minimisation in more than 7 of 48 runs.

The Hamrén and Møller models are similar in stability and are to prefer if that is important for the study.

Runtimes

The complexity, if counted as the number of differential equations, of the models studied is reflected in their runtimes. The Møller model has an advantage over the other models in this aspect, as it has the shortest runtimes. This is most probably due to its simple structure with only 2 differential equations.

Even though the Lledó model was reduced to 13 compartments from the original 25 for the second part of the project it still had the longest runtimes.

Runtimes are highly dependent on initial estimates and how many parameters that needs to be estimated. For consecutive runs the initial estimates were updated, which would have shortened runtimes. Thus one should be careful when interpreting the reported runtimes. As the runs were performed on a computer cluster with nodes of varying processor capacity, the runtime would be highly dependent also on the node processor it was assigned to.

Performance

There are six drug effects investigated in this project. A further analysis could be to look at combinations of these. Also, looking at exactly where to implement the drug effect would perhaps improve the performance in some cases.

All models make use of HbA1c data, but which glucose measurement to use seems to make a great difference in some cases.

The FSI measurements seem to add little information for most drug effects as the Hamrén model mostly outperforms or performs equally good as the de Winter model, though it does not utilise FSI data, even when looking at drug effects linked to insulin, e.g. basal insulin secretion. It should be noted that the MPG-driven models are generally performing better than the FPG-driven, as illustrated in Figure 11c.

Changing to a sparser sampling schedule, such as only weeks 0, 6 and 12, makes the models lose power but more importantly it makes them less stable. Power loss is shown by the difference in Figure 11e and Figure 13e, where especially the de Winter model loses power so that it requires over 600 individuals to get 95% power. The stability loss is seen through the number of runs that crashed, which is higher for all the reduced setups than the standard setting. The de Winter model is most usually the model that need most individuals to reach a certain power, but also the Hamrén model seems to be sensitive to the amount of data provided, see Figure 13a. Putting it differently – reducing the sampling schedule for the Møller and Lledó models can in many cases be done without losing any significant power in the study.

Finding the optimal sampling schedule for a study could save money and time, which is valuable in clinical stages of drug development. It seems that this is highly dependent on which mechanism of effect the drug has and what model is used to describe it. There is a clear trade-off between power and stability.

When looking at the reduced schedule with only the placebo and 25 mg BID arm, it can be seen that more than half the runs crash. This setup only uses 2000 individuals instead of 4000, and since the drug effect was titrated to have a noticeable effect for the highest dose arm, it is not surprising that the power is low in this setup. Nevertheless, the Møller and Hamrén models show stability and good performance while the de Winter and Lledó models crash.

Concluding remarks

We have compared the predictability, stability, runtimes and power to detect drug effect of four previously published models. Overall, using a model where MPG is the driver for HbA1c is superior in stability and power to detect drug effect and the fewer the differential equations in the model the more stable and faster is the analysis. This greatly favours the Møller model. There are however instances where a more mechanistic model may be of greater importance.

Acknowledgements

My deepest thanks to the wonderful people at the Institution of Pharmaceutical Biosciences, especially my supervisor Dr. Maria Kjellsson and Professor Mats Karlsson who has taught me with great passion, knowledge and encouragement. Many thanks to PhD student Steve Choy and Dr. Chenhui Deng for their help in explaining concepts in pharmacokinetics and aiding when I had technical difficulties. I also want to thank my scientific reviewer Dr. Andrew Hooker.

Thanks to Dr. Willem de Winter at University of Leiden and Dr. Jonas Møller at Novo Nordisk for the helpfulness in explaining their models during their respective visits in Uppsala, making this study possible. I would also like to thank PhD student Nicholas Blackburn at University of Tasmania, who helped with feedback on my presentation and is a great friend no matter the distance.

Lastly I would like to thank my mother and my father for being great support.

References

1. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care. 2004 May;27(5):1047–53.

2. American Diabetes Association. Standards of Medical Care in Diabetes--2013. Diabetes Care. 2013 Jan;36(Suppl 1):S11–S66.

3. World Health Organization. Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus – Abbreviated Report of a WHO Consultation. Geneva, World Health Organization, 2011 (WHO/NMH/CHP/CPM/11.1).

4. Kjellsson MC, Cosson VF, Mazer NA, Frey N, Karlsson MO. A Model-Based Approach to Predict Longitudinal HbA1c, Using Early Phase Glucose Data From Type 2 Diabetes Mellitus Patients After Anti-Diabetic Treatment: The Journal of Clinical Pharmacology. J Clin Pharmacol. 2013 Jun;53(6):589–600.

5. Winter W de, DeJongh J, Post T, Ploeger B, Urquhart R, Moules I, et al. 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. 2006 Jun 1;33(3):313–43.

6. Hamrén B, Björk E, Sunzel M, Karlsson MO. Models for Plasma Glucose, HbA1c, and Hemoglobin Interrelationships in Patients with Type 2 Diabetes Following Tesaglitazar Treatment. Clin Pharmacol Ther. 2008 Mar 19;84(2):228–35.

7. Lledó-García R, Mazer NA, Karlsson MO. A semi-mechanistic model of the relationship between average glucose and HbA1c in healthy and diabetic subjects. J Pharmacokinet Pharmacodyn. 2013 Apr 1;40(2):129–42.

8. Møller JB, Overgaard RV, Kjellsson MC, Kristensen NR, Klim S, Ingwersen SH, et al. Longitudinal Modeling of the Relationship Between Mean Plasma Glucose and HbA1c Following Antidiabetic Treatments. Cpt Pharmacometrics Syst Pharmacol. 2013 Oct 30;2(10):e82.

9. Lindbom L, Ribbing J, Jonsson EN. Perl-speaks-NONMEM (PsN)--a Perl module for NONMEM related programming. Comput Methods Programs Biomed. 2004 Aug;75(2):85–94.

10. Lindbom L, Pihlgren P, Jonsson EN, Jonsson N. PsN-Toolkit--a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. Comput Methods Programs Biomed. 2005 Sep;79(3):241–57.

11. Beal S, Sheiner LB, Boeckmann A, & Bauer RJ, NONMEM User's Guides. (1989-2009), Icon Development Solutions, Ellicott City, MD, USA, 2009.

12. R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

13. Jauslin PM, Silber HE, Frey N, Gieschke R, Simonsson USH, Jorga K, et al. An Integrated Glucose-Insulin Model to Describe Oral Glucose Tolerance Test Data in Type 2 Diabetics. J Clin Pharmacol. 2007;47(10):1244–55.

14. Kahn CR, Chen L, Cohen SE. Unraveling the mechanism of action of thiazolidinediones. J Clin Invest. 2000 Dec 1;106(11):1305–7.

15. Vong C, Bergstrand M, Nyberg J, Karlsson MO. Rapid Sample Size Calculations for a Defined Likelihood Ratio Test-Based Power in Mixed-Effects Models. Aaps J. 2012 Feb 17;14(2):176–86.

16. Meininger GE, Scott R, Alba M, Shentu Y, Luo E, Amin H, et al. Effects of MK-0941, a novel glucokinase activator, on glycemic control in insulin-treated patients with type 2 diabetes. Diabetes Care. 2011 Dec;34(12):2560–6.

[i] Reprinted under the Creative Commons license CC BY-NC-ND 3.0 (https://creativecommons.org/licenses/by-nc/3.0/), originally published by Jauslin *et al,* The Journal of Clinical Pharmacology, 2007 (13).

[ii] Reprinted from Springer and Journal of Pharmacokinetics and Pharmacodynamics, 33, 2006, 313-43, 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, Winter W de, DeJongh J, Post T, Ploeger B, Urquhart R, Moules I, *et al*, figure 2, with kind permission from Springer Science and Business Media.

[iii] Reprinted by permission from Macmillan Publishers Ltd: Clinical Pharmacology & Therapeutics, copyright 2008.

[iv] Reprinted from Springer and Journal of Pharmacokinetics and Pharmacodynamics, 40, 2013, 129-42, "A semi-mechanistic model of the relationship between average glucose and HbA1c in healthy and diabetic subjects", Lledó-García R, Mazer NA, Karlsson MO, figure 3, with kind permission from Springer Science and Business Media.

[v] Reprinted under the Creative Commons license CC BY-NC-ND 3.0 (https://creativecommons.org/licenses/by-nc/3.0/), originally published by Møller *et al,* CPT: Pharmacometrics and Systems Pharmacology, 2013 (8). References to Table 2 and Supplementary Material for original article omitted in figure text.