ARTICLE

Study Design Selection in Early Clinical Anti-Hyperglycemic Drug Development: A Simulation Study of Glucose Tolerance Tests

Moustafa M. A. Ibrahim[1,2], Siti M. S. Ghadzi[1,3], Maria C. Kjellsson[1] and Mats O. Karlsson[1]*

In antidiabetic drug development, phase I studies usually involve short-term glucose provocations. Multiple designs are available for these provocations (e.g., meal tolerance tests (MTTs) and graded glucose infusions (GGIs)). With a highly nonlinear, complex system as the glucose homeostasis, the various provocations will contribute with different information offering a rich choice. Here, we investigate the most appropriate study design in phase I for several hypothetical mechanisms of action of a study drug. Five drug effects in diabetes therapeutic areas were investigated using six study designs. Power to detect drug effect was assessed using the likelihood ratio test, whereas precision and accuracy of the quantification of drug effect was assessed using stochastic simulation and estimations. An overall summary was developed to aid designing the studies of antihyperglycemic drug development using model-based analysis. This guidance is to be used when the integrated glucose insulin model is used, involving the investigated drug mechanisms of action.

CPT Pharmacometrics Syst. Pharmacol. (2018) **7**, 432–441; doi:10.1002/psp4.12302; published online on 6 May 2018.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Many designs exist for phase I trials for antihyperglycemic drugs. We know of no systematic evaluation of their relative merits for assessing drug effects of different MoAs.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ We investigated the power, the accuracy, and the precision of pharmacometric analysis of various experimental designs for different drug MoA.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ The most suitable design of phase I antihyperglycemic drug studies depends on and can be predicted from the drug's MoA.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

☑ The phase I trial design choice in the development of antihyperglycemic drugs can be better selected. The exploration through pharmacometric modeling of the relative merits of different trials by drug MoA has the potential to improve drug development in many therapeutic areas.

In the therapeutic area of diabetes, the US Food and Drug Administration guidelines stated that phase I study designs should resemble real life as much as possible.[1] These studies usually involve short-term glucose provocations. With a highly nonlinear, complex system as glucose homeostasis, the various glucose provocations will contribute with different information upon which the probability of detecting the drug effect might be highly dependent. This study focuses mainly on the meal tolerance test (MTT), glucose tolerance test (GTT), repeated fasting blood glucose sampling test (i.e., no provocation (NO) and graded glucose infusion (GGI)). These studies are typically performed after 8–12 hours of overnight fasting. Although these study designs do not allow the assessment of insulin sensitivity, they are much easier to perform and resemble real life more than other study designs, such as glucose clamps.[2] GTTs are performed by giving the patient a standard dose of glucose followed by sampling venous plasma glucose (and optionally plasma insulin) mainly for exploring the clearance of glucose from the blood.[3] The glucose dose is either administered orally (OGTT) or intravenously (IVGTT). Different glucose doses, sampling times, postdose duration of sampling, and analysis methods have been used for GTTs,[4] with 75 g being the recommendation for the oral glucose dose by the World Health Organization. The MTT is the most realistic glucose challenge, in which either a single meal is given with repeated sampling of glucose typically every 30 minutes (i.e., single meal tolerance test (sMTT)),[5] or several meals are given and sampling is performed for 24 hours (i.e., 24-hour meal tolerance test (MTT-24)).[6] Despite being the most real-life mimicking design, MTT-24 is rarely used because of the intensive sampling over the day, whereas sMTT still resembles real-life conditions with the advantage of being less invasive. An alternative to glucose challenges is NO, although it is less informative of the dynamics of glucose, it offers several benefits, such as the low variability and self-monitoring of blood glucose. Graded glucose infusion design comprises five stages of infusion of 20% glucose solution with increasing

[1]Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden; [2]Department of Pharmacy Practice, Helwan University, Cairo, Egypt; [3]School of Pharmaceutical Sciences, Universiti Sains Malaysia, Malaysia. *Correspondence: Mats O. Karlsson (mats.karlsson@farmbio.uu.se)
Received 8 December 2017; accepted 30 March 2018; published online on 6 May 2018. doi:10.1002/psp4.12302

infusion rates.[7] The duration of each stage is typically 40 minutes, in which it starts with a low infusion rate 2 mg/kg and then increases by each stage. It is used to assess glucose-dependent insulin secretion.[8,9] Many targets along the glucose absorption, distribution, metabolism, and excretion scheme are available for development of new antihyperglycemic treatment. The selected mechanisms of action (MoA) for drug effects in this study were stimulation of basal insulin (BINS), glucose clearance, both insulin-dependent glucose clearance (CLG) and independent (CLGI), as well as inhibition of endogenous glucose production (EGP) and absorption of oral glucose (GABS). There are drugs available for treatment of type 2 diabetes, for each of these MoA: increasing insulin secretion by inhibiting K_{ATP} channel of the pancreatic beta cells is the MoA of sulfonylureas[10] and glinides,[11] increasing CLGI by blocking glucose reuptake in renal tubules is the MoA of glycosurics[12] (the US Food and Drug Administration approved canagliflozin in 2013), increasing CLG and decreasing EGP had been introduced as the probable MoA for insulin sensitizers (PPARγ agonists[13] and biguanides,[14] respectively) and decreasing glucose absorption either directly through alpha glucosidase inhibition or by slowing down gastric emptying and promoting satiety is the M oA of acarbose[15] and amylin analogues,[16] respectively.

Pharmacometric model-based analysis had been recently recommended as an alternative for traditional statistical methods to improve clinical trials analysis, particularly after the sharp decline in research and development productivity in the pharmaceutical industry over the past decades.[17] It had been shown that pharmacometric analysis made use of all available data, increasing the overall information content of the trial, and providing an accurate determination of its main aspects as study power and parameter uncertainty.[18] Furthermore, it can be implemented as a decision-making tool in drug development throughout all the different phases of drug development.[19,20]

The aim of this work was to investigate and determine the most appropriate study design in phase I for several hypothetical MoAs of a study drug. Power to detect drug effect, precision, and accuracy of quantification of drug effect were assessed using pharmacometric model-based simulations and estimations.

METHODS

The pharmacometric model

In this work, the integrated glucose insulin (IGI) model developed by Silber *et al.*[21] was used. The model consists of two submodels, glucose and insulin, with interconnecting homeostatic mechanisms. The glucose submodel is a two-compartment model with elimination only from the central compartment. The elimination mechanism is divided into insulin-dependent and insulin-independent, EGP is constant and independent of plasma glucose for patients. The insulin submodel is a one-compartment disposition model, with a first order elimination and endogenous secretion as input. The first phase of insulin secretion is entering the disposition compartment as a system response to administration of bolus glucose but it is absent in patients. There are two feedback mechanisms between the glucose and insulin submodels: the delayed effect of glucose on the second phase of insulin secretion and the delayed effect of insulin on the insulin-dependent glucose elimination. Both delays are mediated through effect compartment models. To fit data following oral glucose administration, the model was modified by Jauslin *et al.*[22] The absorption of glucose following OGTT is described using transit compartments with the addition of incretin effect as an maximum effect (E_{max}) function on insulin second phase secretion. For describing the whole day glucose-insulin profile, including the circadian variation on insulin secretion, a repeated meal model was developed by Jauslin *et al.*[6] These three versions of the IGI model were used to investigate our selected phase I study designs sMTT, MTT-24, OGTT, NO, IVGTT, and GGI.

Study design

The selected glucose provocations were investigated through *in silico* simulations, simulating both glucose and insulin. The main design aspects of each of the standard glucose provocations are listed in **Supplementary Table S1**. The designs were inspired by published studies of glucose experiments.[6,22–24] The MTT-24 design is the only study in which patients received repeated glucose: three meals (8AM, 2PM, and 8PM) and three snacks (11AM, 5PM, and 11PM) and the NO design is the only study design in which no glucose is given. The simulated IVGTT was insulin modified with a 5-minute insulin infusion (0.03 U/kg) 20 minutes after glucose was given.

Drug effects

Drug effects were functions of drug concentrations that were assumed to be at steady-state (Eq. 1). This assumption would correspond to a drug with a long half-life given repeatedly with dosing intervals less than its half-life to minimize fluctuation around its steady-state concentrations:

$$Effect_{drug} = 1 \pm \text{Conc} \cdot \theta_x e^{\eta_{ix}} \quad (1)$$

where θ_x is the estimate of typical slope of the concentration-effect relationship for drug MoA x, Conc is drug steady-state concentration (100 ng/mL), and η_{ix} is the deviation from θ_x for individual *i* and is a normally distributed random variable with zero mean and SD ω_x. As this is a simulation study, the unit of the steady-state concentrations is arbitrarily chosen and will only affect the unit of θ_x. The size of θ_x was determined by titrating the drug effect to produce a 15% decrease in the glucose area under the curve (AUCG) for the typical treated individual compared to the typical untreated individual in the sMTT experiment. The titrated value of θ_x (one for each MoA) was then used in all designs to simulate datasets (one for each design) of 1,000 patients with a crossover design of placebo and drug treatment. The sign of the drug effect determined whether the effect was stimulatory or inhibitory. Five different drug effects were implemented in the model to describe the selected hypothetical drug MoAs, as illustrated in **Figure 1**. For more details on implementation, see **Supplementary model file**, containing the model code for OGTT with all MoAs. As weight was needed for glucose dosing of some designs and in the IGI model for allometric scaling, each patient's weight was randomly simulated assuming a normal distribution, mean 90 $\pm$ 12 kg.[25] The same magnitude

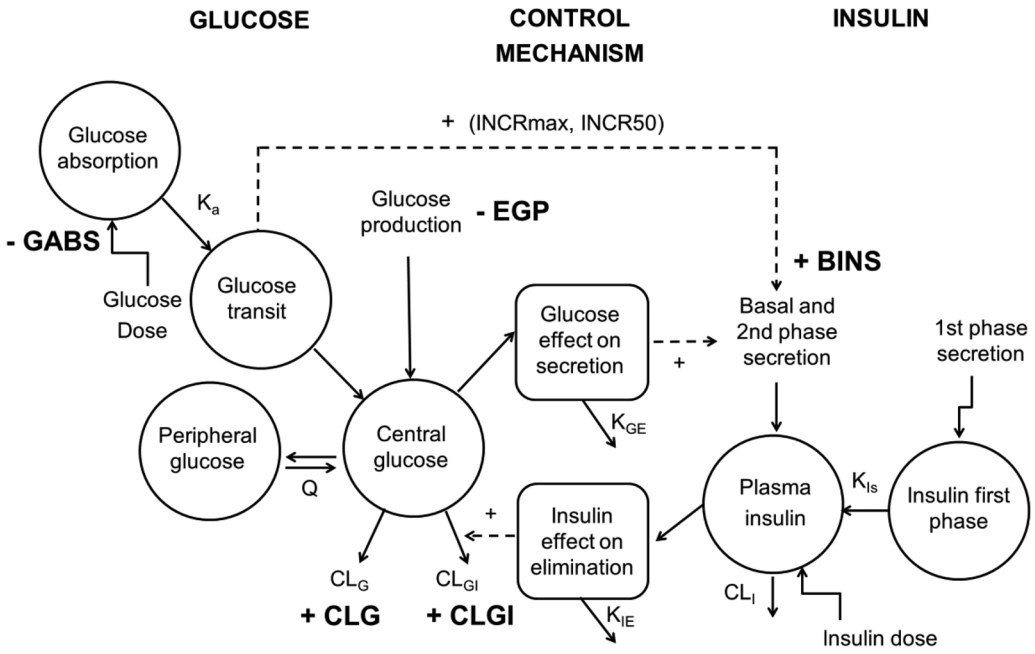

Figure 1 Schematic presentation of the integrated glucose-insulin model for oral tolerance test in type 2 diabetes mellitus[22] with five drug mechanism of actions. BINS, increase basal insulin secretion; CLG, increase insulin-dependent glucose clearance; CLGI, increase insulin-dependent glucose clearance; EGP, decrease glucose production; GABS, decrease glucose absorption; INCR, increase incretin activity.

of θ_x was used in all designs, although titrated based on sMTT, and, thus, the ratio of AUCG between treated and untreated ($\mathrm{AUCGD}/\mathrm{PL_x}$), varied with design and was only for $\mathrm{sMTT} = 15\%$. A washout period of 24 hours was implemented between occasions to re-establish steady-state of the glucose-insulin system. Simulations were based on the published parameter estimates of the IGI models,[6,21,22] and **Figure 2** shows the simulations of a typical individual with and without treatment for each design and drug MoA.

DATA ANALYSIS

Study power calculations

For each of the 6 designs and each of the 5 drug effects, 1 study with 1,000 individuals was simulated, as described above. The model-based power calculations were performed by comparing the difference in objective function value (ΔOFV) between the full model (including drug effect) and the reduced model (without drug effect) fitted to the simulated data (see **Figure 3**). In the estimations, patient-specific and drug-specific parameters were estimated, whereas system-specific parameters were kept fixed; patient-specific parameters in the IGI model are glucose clearance (both insulin-independent and insulin-dependent), glucose and insulin steady-state concentrations, and residual errors for both glucose and insulin, whereas drug-specific parameters were θ_x and ω_x.

Study power for each drug effect was assessed by calculating the relative ratio (RR) of ΔOFV between each study design and sMTT as shown in Eq. 2 where *design* denotes the designs: MTT-24, OGTT, NO, IVGTT, and GGI. This ratio can be interpreted as how many times more individuals are needed for sMTT to achieve equivalent power to the compared design.

$$RR_{design,drug\ \ effect} = \frac{\mathrm{OFV}_{design,\ drug\ \ effect}}{\mathrm{OFV}_{\mathrm{sMTT,\ drug\ \ effect}}} \tag{2}$$

Calculation of accuracy and precision

To assess precision and accuracy of parameter estimation θ_x and AUCG, a stochastic simulation estimation (SSE) was used, in which 500 studies with 15 individuals were simulated, as described in the sections study design and drug effects above, for each of the 6 designs and each of the 5 drug effects (**Figure 3**). The size of the population was based on phase I antihyperglycemic drug clinical trials.[26,27] From these simulations, the true $\mathrm{AUCGD}/\mathrm{PL_{x,true}}$ was calculated. The full model was then fitted to the 500 simulated datasets to produce 500 estimates of $\theta_{x,\mathrm{SSE}}$. The resulting 500 sets of estimates were then used to simulate glucose profiles from which $\mathrm{AUCGD}/\mathrm{PL_{x,SSE}}$ was calculated. Precision and accuracy were assessed by calculating the relative estimation error (REE) of both the parameter θ_x and $\mathrm{AUCGD}/\mathrm{PL_x}$, according to Eqs. 3 and 4.

$$REE_{\theta_x} = \frac{\theta_{x,SSE} - \theta_{x,true}}{\theta_{x,true}} \tag{3}$$

$$REE_{\frac{AUCGD}{PL_x}} = \frac{\frac{AUCGD}{PL_{x,SSE}} - \frac{AUCGD}{PL_{x,true}}}{\frac{AUCGD}{PL_{x,true}}} \tag{4}$$

Power, precision, and accuracy of the studied designs were determined by using the SSE and execute tools implemented in PsN version 3.5,[28] and NONMEM version 7.3.[29] Graphs and dataset creation for NONMEM were performed in R.[30]

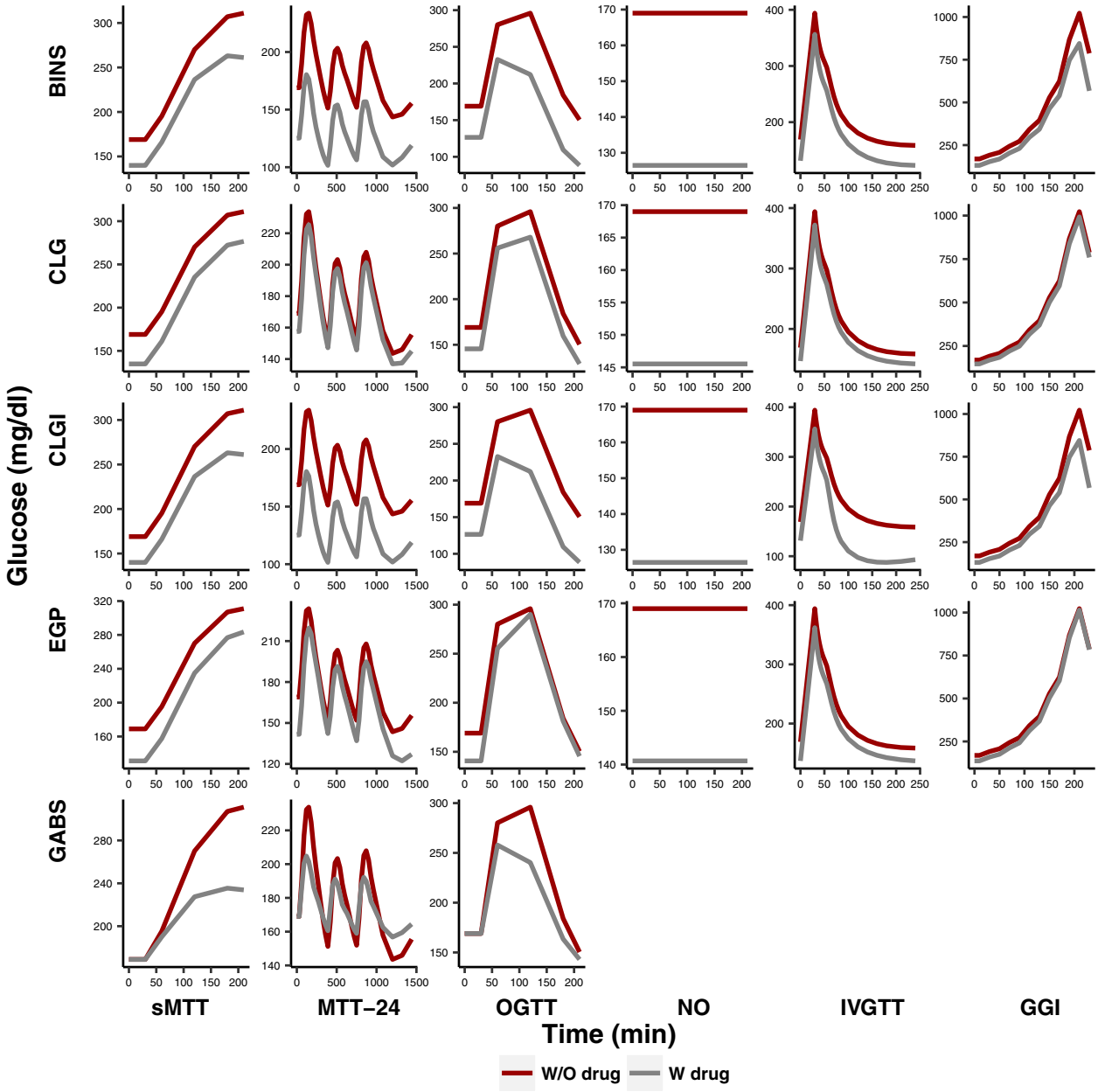

Figure 2 A schematic illustration of the simulated data for a typical individual where the red line indicates the placebo arm and the gray line indicates the drug treatment arm. The different study design models are intravenous glucose tolerance test (IVGTT), oral glucose tolerance test (OGTT), single meal tolerance test (sMTT), 24-hours repeated meal tolerance test (MTT-24), graded glucose infusion (GGI), and repeated fasting samples without provocation (NO). The drug effects are stimulating basal insulin production (BINS), increasing insulin independent glucose clearance (CLG), increasing insulin dependent glucose clearance (CLGI), decreasing endogenous glucose production (EGP), and decreasing oral glucose absorption (GABS). The drug effect is titrated to produce a 15% difference in glucose area under the curve for the sMTT design.

RESULTS

Five drug effects were investigated with six study designs using model-based analysis, as shown in **Figure 2**. The designs NO, IVGTT, and GGI were, for obvious reasons, not investigated in relation to drug effects on glucose absorptions. **Table 1** shows the study power of each study design as the RR to the study power of sMTT model. For all tested MoA, intravenous provocations were always more powerful than oral provocations, with the exception of MTT-24 for certain MoAs. The power of MTT-24 was higher than sMTT for all MoAs, except for increasing CLGI. Repeated fasting sampling NO had for many MoAs higher power than sMTT. Notable though is that safety was not assessed in this study and repeated fasting glucose sampling after drug treatment may be associated with more hypoglycemia than

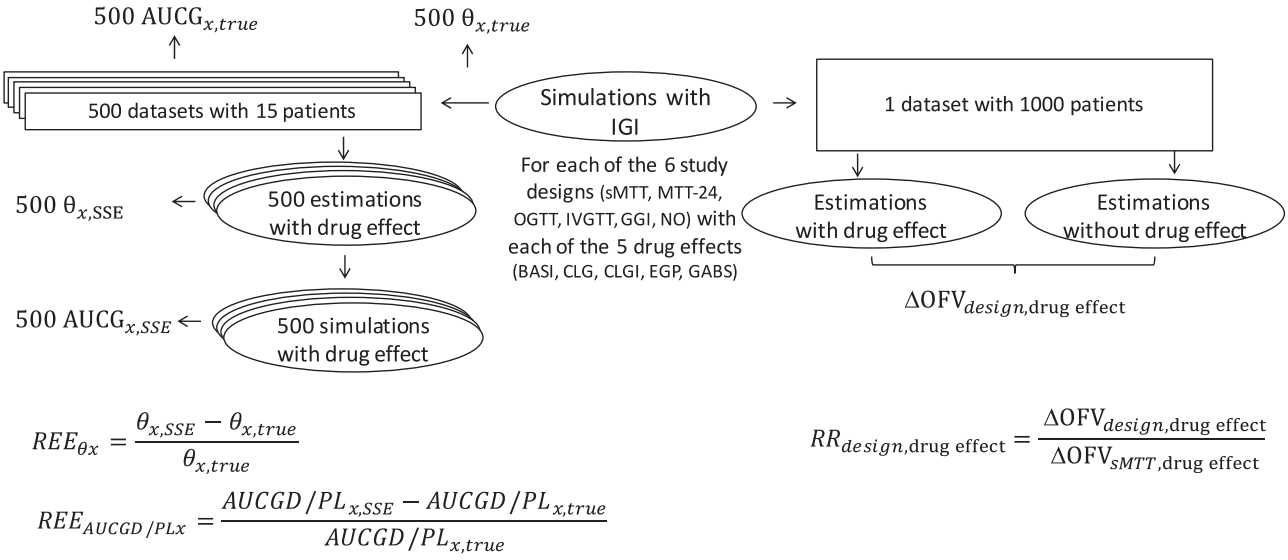

Figure 3 A schematic illustration of study power calculations, in which the appropriate integrated glucose insulin (IGI) model for each study design is used to simulate 1,000 patients with each drug effect, in which after the full and the reduced models of the same study design were fitted to the simulated data, and objective function value (ΔOFV) was calculated for each study design and was used to calculate the relative ratio (RR_{design}). For precision and accuracy calculations, other simulations were performed to produce 500 datasets used to calculate relative estimation error (REE_{θ_x}) and REE_{AUCGD/PLC_x} for each design with each drug effect. AUCG, glucose area under the curve; BASI, basal insulin secretion; EGP, endogenous glucose production; GABS, glucose absorption; GGI, graded glucose infusion; IVGTT, intravenous glucose tolerance test; MTT-24, 24-hour meal tolerance test; NO, no provocation; OGTT, oral glucose tolerance test; sMTT, single meal tolerance test; SSE, stochastic simulation and estimation.

Table 1 The study power of each design given as the ratio for how many times more individuals are needed for **sMTT** to achieve equivalent power to the compared design

		Study designs					
		sMTT	**MTT-24**	**OGTT**	**NO**	**IVGTT**	**GGI**
Drug effects	BINS	1	8.7	1.5	2.0	5.8	5.4
	Increasing CLG	1	0.6	0.3	0.8	2.8	3.9
	Increasing CLGI	1	17.3	2.8	2.7	18.0	7.9
	Decreasing EGP	1	2.0	0.6	1.7	4.2	4.8
	Decreasing GABS	1	3.9	0.4	–	–	–

The red color represents the most powerful study design and the gray represents the second most powerful study design for each mechanism of action.
BINS, stimulating basal insulin production; CLG, insulin-independent glucose clearance; CLGI, insulin-dependent glucose clearance; EGP, endogenous glucose production; GABS, glucose absorption; GGI, graded glucose infusion; IVGTT, intravenous glucose tolerance test; NO, no provocations test; OGTT, oral glucose tolerance test; sMTT, single meal tolerance test; MTT-24, 24-hour meal tolerance test.

other designs. For drug effects increasing the CLGI, the MTT-24 and OGTT were the least powerful study designs. Potentially, the stimulation of insulin release, at a higher rate by incretin activity, made CLGI less important.

Precision and accuracy of model parameters and simulated glucose concentrations were computed by SSE. REE_{θ_x} and REE_{AUCGD/PL_x} are visualized in **Figure 4** and **Figure 5**, respectively. The precision and accuracy of $AUCGD/PL_x$ was, in general, better than for the θ_x. All investigated designs for drug effects on basal insulin and glucose absorption performed well in terms of accuracy and precision, thus, for these MoAs, the power to detect a drug effect is the main difference. Those designs with a good precision and accuracy of θ_x performed well in terms of glucose simulations also. Only for drug effects on CLG with NO, was the bias large in both θ_{CLGI} and $AUCGD/PL_{CLGI}$. There was also a small, although significant bias for both θ_{CLG} and $AUCGD/PL_{CLG}$ for NO. Thus, although the NO design is surprisingly powerful, it resulted in the worst precision and accuracy. The accuracy and precision of parameters and glucose concentrations derived from OGTT, IVGTT, and GGI were overall good for all MoAs. Thus, the most powerful study designs, albeit including the two most invasive, are also producing accurate and precise estimates. Both the sMTT and the MTT-24 designs were disappointingly biased in terms of θ_x. These designs mimic real life the most, however, the accuracy was poor for θ_{CLG}, θ_{CLGI}, and θ_{EGP}, and, in particular, the precision for sMTT of θ_{CLGI}. Slightly reassuring was the results of accuracy and precision of $AUCGD/PL_x$ for these designs, as they performed similar to other designs.

In summary, the study design with the highest power for a drug effect was not always the most precise neither was it the most accurate. The overall most powerful, accurate, and precise study designs were the intravenous glucose

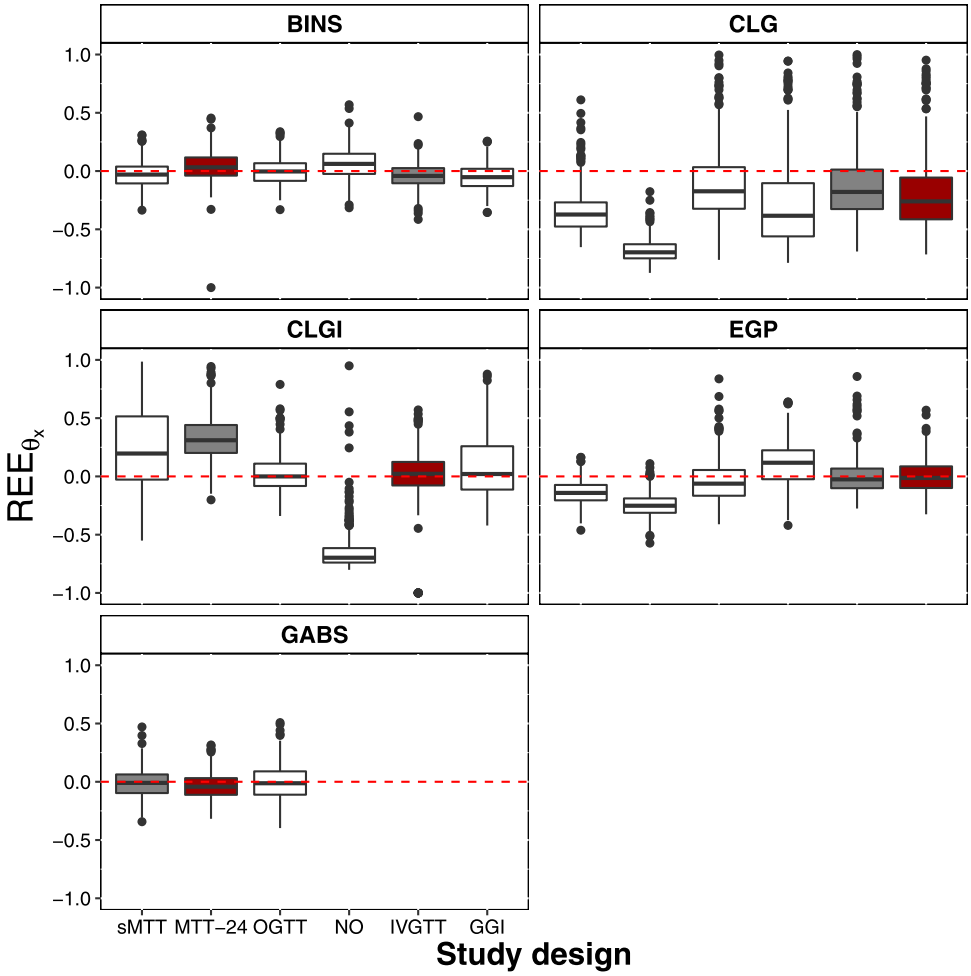

Figure 4 A descriptive representation of the distribution of relative estimation error (REE$_{\theta_x}$) of each of the study designs for all of the drug effects. The different study designs are intravenous glucose tolerance test (IVGTT), oral glucose tolerance test (OGTT), single meal tolerance test (sMTT), 24-hour meal tolerance test (MTT-24), graded glucose infusion (GGI), and no provocations test (NO). The average convergence rate for all estimations in the SSE was 0.713. BINS, increase basal insulin secretion; CLG, increase insulin-dependent glucose clearance; CLGI, increase insulin-independent glucose clearance; EGP, decrease glucose production; GABS, decrease glucose absorption.

administration experiments. From a power perspective, MTT-24 seems like a good alternative, apart from effects on CLGI (e.g., SGLT2-inhibitors).

DISCUSSION

In phase I antihyperglycemic drug studies, various study designs carry different information according to their respective ways in provoking glucose homeostasis systems. Thus, few therapeutic areas have so many designs from which to choose as diabetes. In this work, we have tried to outline what the relative merits are of different designs in terms of identifying and estimating drug effects and how their relative rank depends on the MoA.

The sMTT shows lower and slower glucose profiles than intravenous provocations as intestinal glucose and lipids increase insulin response through incretin peptides. Although the MTT-24 provides similar glucose and insulin profiles to sMTT, with more samples, longer study duration, and also nightly suppression of the insulin production. For these two designs, the MTT version of the IGI model was used for data creation and analysis, this model contains a linear incretin effect. The OGTT shows similar glucose profiles to sMTT but with faster absorption phase and clearer elimination phase, as a consequence of incretin effects, second phase response of insulin to glucose is much higher than intravenous, however, incretin response of the OGTT is lower than sMTT due to the lack of intestinal lipids in OGTT. The OGTT version of the IGI model was used to create and analyze data for this provocation and this model includes an E_{max} model for the incretin effect. Repeated fasting samples NO is expected to be the most informative for drug effects acting on mechanisms affecting the glucose and insulin steady states, these data were produced and analyzed using the IGI model for OGTT without intestinal glucose-related incretin effect. Because the full dose enters the body instantly during IVGTT, a stronger insulin secretion

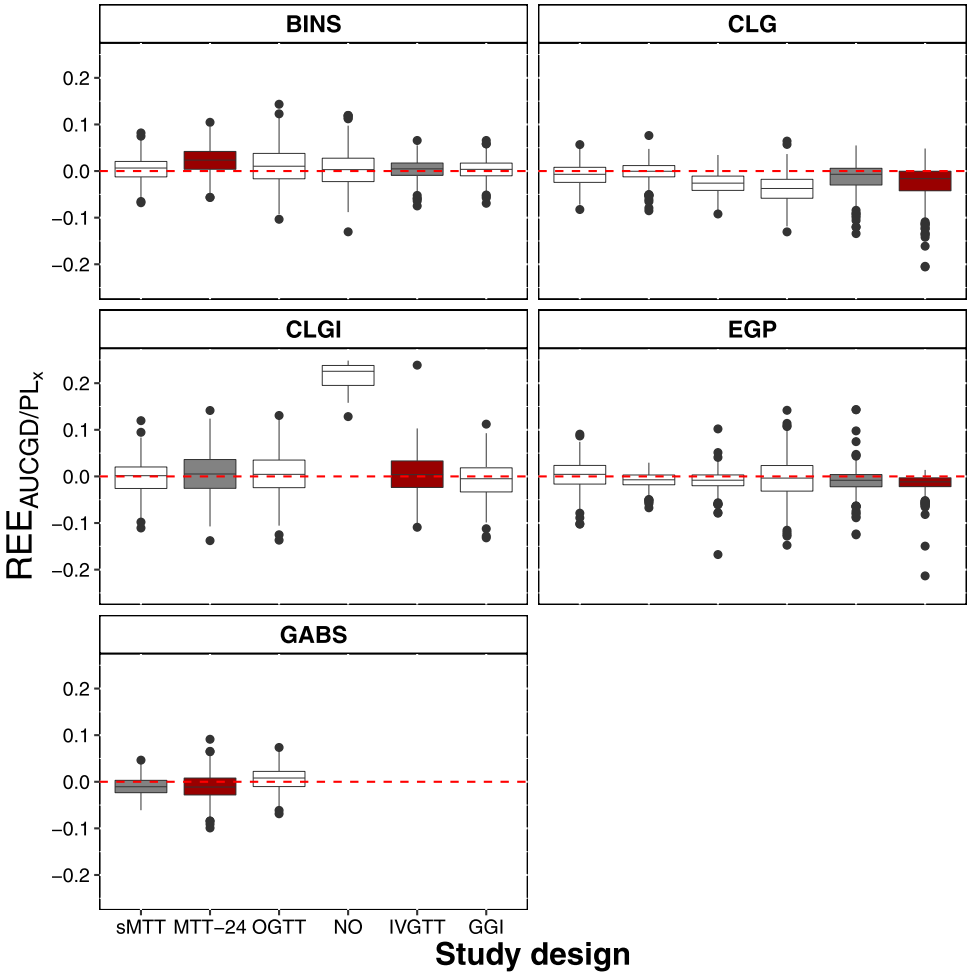

Figure 5 A descriptive representation of the distribution of relative estimation error (REE_{AUCGD/PL_x}) of each of the study designs for all drug effects. The different study design models are intravenous glucose tolerance test (IVGTT), oral glucose tolerance test (OGTT), single meal tolerance test (sMTT), 24-hour meal tolerance test (MTT-24), graded glucose infusion (GGI), and no provocations test (NO). The red color indicates the most powerful study design and the gray indicates the second most powerful study design. The average convergence rate for all estimations in the stochastic simulation estimation was 0.713. AUCG, glucose area under the curve; BINS, increase basal insulin secretion; CLG, increase insulin-independent glucose clearance; CLGI, increase insulin-dependent glucose clearance; EGP, decrease glucose production; GABS, decrease glucose absorption.

with shorter duration is observed, compared with oral glucose provocations. The GGI produces a stepwise increase in plasma glucose levels that allows the assessment of glucose dependent insulin secretion as well as glucose elimination. The IVGTT version of the IGI model was used to create and analyze data for these two provocations and this model includes no incretin effect. These various ways in provoking the glucose homeostasis explain the differences seen in the final parameter estimates of the different versions of I GI model, in particular for the estimates of glucose clearance. It has been shown that the IGI model can correctly describe data from different experiments and real-life situations.[12] Reality imitation with the simulated data for various drug effects shown in **Figure 2** is an example to the applicability of the IGI model in understanding antihyperglycemic drug actions.

The results of the power analysis can largely be explained by the difference in sampling schedule between the designs. The designs differed greatly in position and number of observations per subject ($n = 37$, 26, 14, and 6 for MTT-24, IVGTT, GGI, and OGTT/sMTT/NO, respectively). Thus, much of the higher power observed for IVGTT, GGI, and MTT-24 is attributed to the sampling schedules. To investigate the impact of the sampling schedule, the RR was calculated for NO and O GTT with the same sampling schedule as the IVGTT design. This resulted in a massive change in RR for both NO and OG TT (see **Supplementary Table S2**), with NO now being the most powerful of all designs for all drug effects, except CLG.

Although the sampling schedule explains most of the results, the power is also related to how glucose and insulin responses are provoked in the design. The most powerful design to detect drug effects on glucose absorption was the MTT-24, mainly due to the large number of observations but also the repeated oral intake of glucose. As glucose is absorbed several times, the power to detect a drug effect on glucose absorption should be high. When comparing OGTT

Table 2 Recommendation on study design based on REE_{θ_x}, REE_{AUCGD/PL_x}, and hypothesized site of drug mechanism of actions

	REE_{θ_x}		REE_{AUCGD/PL_x}	
Drug effects	**Best**	**Less invasive alternative**	**Best**	**Less invasive alternative**
BINS	MTT-24	NO	IVGTT	NO
CLG	GGI	sMTT	MTT-24	sMTT
CLGI	IVGTT	OGTT	IVGTT	OGTT
EGP	GGI	NO	GGI	NO
GABS	MTT-24	sMTT	MTT-24	sMTT

AUCGD, glucose area under the curve; BINS, increase basal insulin secretion; CLG, increase insulin-independent glucose clearance; CLGI, increase insulin-dependent glucose clearance; EGP, decrease glucose production; GABS, decrease glucose absorption; GGI, graded glucose infusion; IVGTT, intravenous glucose tolerance test; MTT-24, 24-hours meal tolerance test; NO, no provocations test; OGTT, oral glucose tolerance test; REE, relative estimation error; sMTT, single meal tolerance test.

to sMTT, with the same number of observations, the sMTT is more powerful, this can be explained by the slower absorption rate in the sMTT design, with a longer duration allowing for easier detection of changes in absorption process.

Hypothetically, when insulin concentrations are high, the relative importance of insulin-dependent glucose elimination will be high compared to basal EGP, basal insulin secretion, and insulin-independent glucose elimination. Thus, designs provoking a strong insulin response will have a lower power to detect drug effects on basal glucose production, basal insulin secretion, and CLGI, whereas the opposite applies for drug effects on CLG. Another way of explaining it is to look at the change in fasting plasma glucose (FPG). Drug effects with their main effect on FPG will have a much higher signal-to-noise ratio in designs in which glucose is invariable, such as in the NO design. This is true for basal EGP, basal insulin secretion, and CLGI, however, not for CLG.

If we first examine drug effects increasing basal insulin secretion, this drug effect reduces fasting glucose concentration from 170 mg/dL to ~140 mg/dL. When accounting for the difference in the number of observations of the designs, the highest power was observed for NO, which is in line with this hypothesis.

For drug effects on EGP, there is a similar trend; the reduction in fasting glucose concentrations is larger, ~40 mg/dL. Thus, NO will be the most powerful study design for EGP when accounting for the number of observations. In addition, the order of power does not follow the order of number of observations, indicating that the number of observations for this MoA is of less importance. The order of power between the frequently sampled designs MTT-24, GGI, and IVGTT, is driven by which of the designs have the smallest provoked glucose-insulin response, which is GGI.

For drug effects on CLGI, NO and OGTT have the highest power. It seems as if sMTT is a fairly powerful design for this MoA, given that the RRs for all designs for this MoA are the smallest.

Last, the power to detect a drug effect on CLG should, unlike the power for other drug effects, benefit from provoking insulin response. When accounting for the uneven number of observations, NO and OGTT are powerful but not as powerful as MTT-24 or IVGTT, which is the design with the highest insulin concentrations and should be the most powerful.

In our investigations, RRs were used. As a consequence, any changes in absolute power that are the same for the designs will not affect the relative power as they cancel out (i.e., magnitude of drug effect and variability; e.g., interindividual and interoccasion variability as well as pharmacokinetic and pharmacodynamic variability), linear or nonlinear pharmacokinetic-pharmacodynamic relationship. In addition, the results are independent of which statistical test was used as long as the metric uses difference in OFV between nested models (e.g., the likelihood ratio test or the Akaike Information Criterion).

In terms of precision, it has been shown that despite the biased parameter estimates of some models, they were capable of precisely and accurately simulating plasma AUCG. This reflects a good model performance in terms of simulating and estimating the $AUCGD/PL_x$. Thus, predicting the size of the drug effect in terms of AUCG is possible using parameter estimates from most of the designs. However, caution should be taken when interpreting the actual parameter values and statements like "the EGP increased with 15%" should only be considered for those MoAs and designs in which the parameter estimates were not biased and sufficiently precise.

Due to run-time issues, we calculated the fraction of the needed number of individuals for the sMTT design to achieve the same power as other studied designs, instead of the actual needed number of subjects to achieve a particular study power. To do the latter, the true type I error rate would have to be assessed, which can be done using the simulation hypothesis test. Once the type I error rate is known, a full SSE or Monte-Carlo Mapped Power Method,[31] can be used to assess the number of individuals required for a certain power, both simulation hypothesis test and SSE rely on repeated dataset simulations and estimations, which is unrealistic with a model, such as the IGI model, in which the run-times are long. An alternative method, used in this work, when comparing study power, is to calculate the ratio between the ΔOFVs of the compared study designs. This method does not require assessment of the type I error rate or repeated sampling and is thus faster than the above alternatives. This method has been shown to produce similar results to the ratio of individuals at a certain power using the Monte-Carlo Mapped Power Method.[32] Yet, another way to compare these designs

would be to compare the SE of θ_x estimates, in a similar manner to efficiency calculations in optimal design. With optimal design, it is possible also to change the standard design variables to increase the power for a certain drug effect.[33] However, this was not investigated in the current work.

Based on the results from this study, an overall summary was developed to aid designing studies of antihyperglycemic drug development using model-based analysis. This guidance can be used in a scenario when the IGI model is used, involving a certain expected drug MoA. The best study design for different analysis purposes and its less invasive alternatives for each hypothetical drug MoA are summarized in **Table 2**. The results presented in this article represent a model-based analysis. Although not investigated, it is possible that the results would hold true for a more traditional analysis, such as a *t*-test of AUCG with and without treatment; that remains, however, to be verified. Notable though is that the results would be invalidated if a *t*-test of AUCG was performed on baseline-corrected AUCG as the information of drug effects on FPG would be ignored.

CONCLUSIONS

The selection of study design for antidiabetic phase I studies may be complex and is commonly decided by patient convenience or study feasibility. The work presented here provides an insight in the relative merits of the studied designs when performing pharmacometric analysis and will allow more informed selection among the designs. This work also shows that pharmacometric simulation is a valuable tool in the design of phase I studies in antihyperglycemic drug development in which the power, precision, and accuracy of various study designs is dependent on the MoA of the study drug.

Source of Funding. The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115156, resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007–2013) and EFPIA companies' in kind contribution. The DDMoRe project is also financially supported by contributions from academic and SME partners.

Conflict of Interest. The authors declared no competing interests for this work.

Author Contributions. M.M.A.I., S.M.S.G., M.C.K., and M.O.K. wrote the manuscript. M.M.A.I., S.M.S.G., M.C.K., and M.O.K. designed the research. M.M.A.I. performed the research. M.M.A.I., S.M.S.G., M.C.K., and M.O.K. analyzed the data.

1. U.S. Department of Health and Human, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Guidance for industry. Diabetes mellitus: developing drugs and therapeutic biologics for treatment and prevention. Clinical/Medical. 2008;(February):1–30. <http://www.fda.gov/downloads/Drugs/Guidances/ucm071624.pdf>.
2. De Fronzo, R.A., Tobin, J.D. & Andres, R. Glucose clamp technique: a method for quantifying insulin secretion and resistance. *Am. J. Physiol.* **237**, E214–E223 (1979).
3. Conn, J.W. Interpretation of the glucose tolerance test. The necessity of a standard preparatory diet. *Am. J. Med. Sci.* **199**, 555–564 (1940).
4. Institute for Quality and Efficiency in Health Care. Glucose tolerance test: how does it work exactly? Informed Health Online. <https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0072515/>. Accessed 22 June 2013.
5. Dalla Man, C., Piccinini, F., Basu, R., Basu, A., Rizza, R.A. & Cobelli, C. Modeling hepatic insulin sensitivity during a meal: validation against the euglycemic hyperinsulinemic clamp. *AJP Endocrinol. Metab.* **304**, E819–E825 (2013).
6. Jauslin, P.M., Frey, N. & Karlsson, M.O. Modeling of 24-hour glucose and insulin profiles of patients with type 2 diabetes. *J. Clin. Pharmacol.* **51**, 153–164 (2011).
7. Brandt, A., Katschinski, M., Arnold, R., Polonsky, K.S., Goke, B. & Byrne, M.M. GLP-1-induced alterations in the glucose-stimulated insulin secretory dose-response curve. *Am. J. Physiol.* **281**, E242–E247 (2001).
8. Byrne, M.M., Sturis, J. & Polonsky, K.S. Insulin secretion and clearance during low-dose graded glucose infusion. *Am. J. Physiol.* **268**(1 Pt 1), E21–E27 (1995).
9. Byrne, M.M., Sturis, J., Sobel, R.J. & Polonsky, K.S. Elevated plasma glucose 2 h post challenge predicts defects in beta-cell function. *Am. J. Physiol.* **270**(4 Pt 1), E572–E579 (1996).
10. Nguyen, N.D.T. & Le, L.T. Targeted proteins for diabetes drug design. *Adv. Nat. Sci. Nanosci. Nanotech.* **3**, 013001 (2012).
11. Culy, C.R. & Jarvis, B. Repaglinide: a review of its therapeutic use in type 2 diabetes mellitus. *Drugs* **61**, 1625–1660 (2001).
12. Chao, E.C. & Henry, R.R. SGLT2 inhibition–a novel strategy for diabetes treatment. *Nat. Rev. Drug Discov.* **9**, 551–559 (2010).
13. Berger, J. & Moller, D.E. The mechanisms of action of PPARs. *Annu. Rev. Med.* **53**, 409–435 (2002).
14. Miller, R.A., Chu, Q., Xie, J., Foretz, M., Viollet, B. & Birnbaum, M.J. Biguanides suppress hepatic glucagon signalling by decreasing production of cyclic AMP. *Nature* **494**, 256–260 (2013).
15. Van de Laar, F.A., Lucassen, P.L., Akkermans, R.P., Van de Lisdonk, E.H., Rutten, G.E. & Van Weel, C. Alpha-glucosidase inhibitors for type 2 diabetes mellitus. *Cochrane Database Syst. Rev.* **2**, CD003639 (2005).
16. Jones, M.C. Therapies for diabetes: pramlintide and exenatide. *Am. Fam. Physician* **75**, 1831–1835 (2007).
17. Kola, I. & Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* **3**, 711–715 (2004).
18. Jonsson, E.N. & Sheiner, L.B. More efficient clinical trials through use of scientific model-based statistical tests. *Clin. Pharmacol. Ther.* **72**, 603–614 (2002).
19. Lockwood, P., Ewy, W., Hermann, D. & Holford, N. Application of clinical trial simulation to compare proof-of-concept study designs for drugs with a slow onset of effect; an example in Alzheimer's disease. *Pharm. Res.* **23**, 2050–2059 (2006).
20. Karlsson, K.E., Vong, C., Bergstrand, M., Jonsson, E.N. & Karlsson, M.O. Comparisons of analysis methods for proof-of-concept trials. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e23 (2013).
21. Silber, H.E., Jauslin, P.M, Frey, N., Gieschke, R., Simonsson, U.S. & Karlsson, M.O. An integrated model for glucose and insulin regulation in healthy volunteers and type 2 diabetic patients following intravenous glucose provocations. *J. Clin. Pharmacol.* **47**, 1159–1171 (2007).
22. Jauslin, P.M. *et al.* An integrated glucose-insulin model to describe oral glucose tolerance test data in type 2 diabetics. *J. Clin. Pharmacol.* **47**, 1244–1255 (2007).
23. Jauslin, P.M., Karlsson, M.O. & Frey, N. Identification of the mechanism of action of a glucokinase activator from oral glucose tolerance test data in type 2 diabetic patients based on an integrated glucose-insulin model. *J. Clin. Pharmacol.* **52**, 1861–1871 (2012).
24. Wang, X. *et al.* Quantification of beta-cell insulin secretory function using a graded glucose infusion with C-peptide deconvolution in dysmetabolic, and diabetic cynomolgus monkeys. *Diabetol. Metab. Syndr.* **5**, 40 (2013).
25. Kjellsson, M.C., Cosson, V.F., Mazer, N.A., Frey, N. & Karlsson, M.O. A model-based approach to predict longitudinal hba1c, using early phase glucose data from type 2 diabetes mellitus patients after anti-diabetic treatment. *J. Clin. Pharmacol.* **53**, 589–600 (2013).
26. Røge, R.M., Klim, S., Kristensen, N.R., Ingwersen, S.H. & Kjellsson, M.C. Modeling of 24-hour glucose and insulin profiles in patients with type 2 diabetes mellitus treated with biphasic insulin aspart. *J. Clin. Pharmacol.* **54**, 809–817 (2014).
27. Røge, R.M., Klim, S., Ingwersen, S.H., Kjellsson, M.C. & Kristensen, N.R. The effects of a GLP-1 analog on glucose homeostasis in type 2 diabetes mellitus quantified by an integrated glucose insulin model. *CPT Pharmacometrics Syst. Pharmacol.* **4**, 1–9 (2015).
28. Lindbom, L., Pihlgren, P. & Jonsson, E.N. PsN-toolkit—a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput. Methods Programs Biomed.* **79**, 241–257 (2005).
29. Beal, S., Sheiner, L.B., Boeckmann, A. & Bauer, R.J. NONMEM User's Guides (1989–2009). Ellicott City, MD: Icon Development Solutions; 2009.
30. Team RC. R: A Language and Environment for Statistical Computing. Vienna, Austria: 2014. <http://www.R-project.org>.

31. Vong, C., Bergstrand, M., Nyberg, J. & Karlsson, M.O. Rapid sample size calculations for a defined likelihood ratio test-based power in mixed-effects models. *AAPS J.* **14**, 176–186 (2012).
32. Sheikh Ghadzi, S.M., Karlsson, M.O. & Kjellsson, M.C. Implications for drug characterization in glucose tolerance tests without insulin: simulation study of power and predictions using model-based analysis. *CPT Pharmacometrics Syst. Pharmacol.* **6**, 686–694 (2017).
33. Ueckert, S., Hennig, S., Nyberg, J., Karlsson, M.O. & Hooker, A.C. Optimizing disease progression study designs for drug effect discrimination. *J. Pharmacokinet. Pharmacodyn.* **40**, 587–596 (2013).

Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (http://psp-journal.com)