

ARTICLE

Comparison of Power, Prognosis, and Extrapolation Properties of Four Population Pharmacodynamic Models of HbA1c for Type 2 Diabetes

Gustaf J. Wellhagen, Mats O. Karlsson and Maria C. Kjellsson*

Reusing published models saves time; time to be used for informing decisions in drug development. In antihyperglycemic drug development, several published HbA1c models are available but selecting the appropriate model for a particular purpose is challenging. This study aims at helping selection by investigating four HbA1c models, specifically the ability to identify drug effects (shape, site of action, and power) and simulation properties. All models could identify glucose effect nonlinearities, although for detecting the site of action, a mechanistic glucose model was needed. Power was highest for models using mean plasma glucose to drive HbA1c formation. Insulin contribution to power varied greatly depending on the drug target; it was beneficial only if the drug target was insulin secretion. All investigated models showed good simulation properties. However, extrapolation with the mechanistic model beyond 12 weeks resulted in drug effect overprediction. This investigation aids drug development in decisions regarding model choice if reusing published HbA1c models.

CPT Pharmacometrics Syst. Pharmacol. (2018) 7, 331–341; doi:10.1002/psp4.12290; published online 25 March 2018.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Several HbA1c models are available in literature to be reused, with some aspects being different and others similar.

WHAT QUESTION DID THIS STUDY ADDRESS?

This study was performed to guide selection of the appropriate HbA1c model.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

Model choice will depend on the drug mechanism of action as well as the main analysis purpose. All models identified nonlinearity in glucose; however, a mechanistic model would differentiate the accurate site of drug

action. Power of detecting drug effects was marginally higher for models using MPG, instead of FPG for HbA1c formation, unless the drug effect was incretin. In addition, insulin measurements were only beneficial if insulin was the main target. All tested models showed good prognostic and extrapolative properties, except the de Winter *et al.*¹² model, which overpredicted the drug effect for 26-week extrapolation.

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

Modelers can reuse published models that are fit for purpose, knowing their strengths and weaknesses. Focus can instead be put on interpreting the results and making informed decisions in drug development.

Reusing published models is attractive as it saves time; time to be used for informing decisions in drug development and increasing certainty of the decision. Diabetes is an area in which many models are available, describing mainly glycosylated hemoglobin (HbA1c).¹ HbA1c (measuring sustained glycemic control) is commonly used as the main biomarker in antihyperglycemic drug development, along with fasting plasma glucose (FPG) and/or fasting serum insulin (FSI), and possibly mean plasma glucose (MPG). In phase II of antihyperglycemic drug development, study designs are often similar: double-blinded, parallel, and placebo-controlled, with 4–5 longitudinal observations of biomarkers over 12–14 weeks, in which the study drug is given as an add-on treatment to the standard-of-care (e.g., metformin).² Modeling is well-accepted in diabetes research, thus, finding published models to reuse for pharmacometric analysis of phase II study data are easy and the challenge rather lays in selecting a model fit for purpose.

The purpose of a pharmacometric analysis varies greatly: selection of optimal doses and dosing regimens to maximize the HbA1c-lowering effect in phase II, quantitative support for the decision of dose reduction in subsequent phase II trials,³ mechanistic understanding of drug effects on hemoglobin (Hb) and its impact on HbA1c,⁴ assessment of demographics and disease progression on HbA1c to inform patient inclusion criteria,⁵ evaluation of the impact of genetic pharmacokinetic differences on HbA1c,⁶ investigation of differences in HbA1c effects between once-daily and twice-daily dosing,⁷ and all prospective predictions of HbA1c in phase II using phase I short-term glucose,⁸ phase III using phase II study data,⁹ and patients using data from healthy subjects.¹⁰ In most of these examples, pharmacometric analysis was also used to quantify uncertainty.

Some aspects of the HbA1c models are shared and yet others are different. To guide modelers in the choice

of the models for reuse, we herein present a quantitative, simulation-based investigation of four already published HbA1c models, demonstrating their respective strengths and weaknesses. Four aspects were compared: (1) the ability to identify the correct mechanism of action; (2) the power to detect a drug effect; (3) simulation performance for similar study duration (prognosis); and (4) longer studies (extrapolation).

METHODS

Models

The four HbA1c models investigated throughout this article are denoted: A Dynamic HbA1c EndpOint Prediction Tool (ADOPT) published by Møller *et al.*,¹¹ 2013; FPG-FSI-HbA1c (FFH) by de Winter *et al.*,¹² 2006; FPG-Hb-HbA1c (FHH) by Hamrén *et al.*,⁴ 2008; and Integrated Glucose-RBC-HbA1c (IGRH) by Lledó-García *et al.*,¹³ 2013. Schematic pictures of all HbA1c models are shown in **Figure 1**.^{4,11–13} They represent different complexity, ranging from indirect response to more complex mechanistic models, utilizing different biomarkers. The selected models have all been used to guide decisions in drug development.

ADOPT

The ADOPT model (**Figure 1a**) consists of two indirect response models: one for MPG and one for HbA1c. The HbA1c formation is determined by MPG with a first order rate.

FFH

The FFH model (**Figure 1b**) contains three linked indirect response models describing FSI, FPG, and HbA1c. It was developed to describe the type 2 diabetes disease progression with terms of beta cell function (EFB) and insulin sensitivity (EFS). The model contains a mechanistic homeostatic feedback between FPG and FSI, in which FPG affects the production rate of FSI and FSI affects the production rate of FPG. The relationship is described using the homeostasis model of assessment (HOMA) equations,¹⁴ in which healthy subjects have a HOMA-%B = 1 and HOMA-%S = 1 at FPG = 4.5 mM and FSI = 5 mU/L, which at steady-state gives $B_0 = FSI_0 / (5 * (FPG_0 - 3.5)) = HOMA\text{-}\%B$ and $S_0 = FPG_0 * FSI_0 / 22.5 = HOMA\text{-}\%S$. The HbA1c is formed by an FPG-driven first order rate and a non-FPG-driven constant rate representing glycosylation from postprandial glucose (PPG) and contamination in HbA1c, unrelated to glucose, respectively.

FHH

The FHH model (**Figure 1c**) describes HbA1c formation through the life-span of red blood cells (RBCs) with eight transit compartments: four nonglycosylated and four glycosylated compartments of RBC of varying age. During their lifetime, RBCs are glycosylated with a first order rate driven by $f(FPG)$. The FPG is modeled with an indirect response model and $f(FPG)$ is a power function of FPG, allowing for a nonlinear relationship between FPG and HbA1c formation. This nonlinearity has been hypothesized to be related to FPG only measuring fasting glucose exposure, excluding additional exposure of PPG.

IGRH

The IGRH model (**Figure 1d**) is similar to FHH with HbA1c formation described through the RBC life-span, however, using more transit compartments. In this model, MPG drives the glycosylation in a linear first order process. The glucose exposure affects the life-span of RBCs, so that higher glucose levels shorten the life-span of RBCs. Precursors of RBCs are also assumed to be partly glycosylated. MPG is modeled with an indirect response model, similar to the ADOPT model. The total number of compartments for RBC in the IGRH model was 24 in the original publication and for simulating data in this study. However, for reduced estimation time, 12 transit compartments were used in estimations: 6 for nonglycosylated and 6 for glycosylated RBCs. Reducing the number of compartments has little effect on the fit.¹⁵

Simulation study design

Data were simulated according to a 12-week parallel group phase IIa clinical trial design with patients with type 2 diabetes. Inclusion criteria on baseline were HbA1c 7–10%, FPG 7–13.3 mM, and FSI 7.3–18 mU/L. A total of 4,000 patients were divided equally in four treatment arms: (1) the placebo arm; (2) the 10 mg b.i.d. arm; (3) the 25 mg b.i.d. arm; and (4) the 50 mg b.i.d. arm. Patients had a standardized diet of three meals and two snacks per day, corresponding to 62.5 g and 12.5 g glucose, respectively.

Glucose and insulin were simulated with the integrated glucose-insulin (IGI) model,¹⁶ shown in **Supplementary Figure S1**, with parameters fixed to published estimates. The glucose and insulin error was set to 10.7% and 27%, respectively, to include both assay error and model misspecification.^{17,18} Placebo was described as a -0.1 ± 0.14 mM (mean \pm SD) change in glucose at steady-state; a magnitude in line with a number of placebo-controlled antihyperglycemic drug trials.^{19–27}

Five different drug effects were investigated separately, each incorporated at a different site in the IGI model and drugs were assumed to elicit their effect without a delay: stimulation of basal insulin secretion (BASI) similar to the expected effect of an insulin secretagogue, insulin-independent glucose (CLG) and insulin-dependent (CLGI) glucose elimination mimicking the effects of SLGT-2 inhibitors and peroxisome proliferator-activated receptor agonists, respectively, inhibition of endogenous glucose production (EGP), hypothesized as the main action of metformin and stimulation of incretin release (INCR), similar to the expected effect of dipeptidyl peptidase-4 inhibitors. **Supplementary Equation S1** details how the drug effects were implemented in the simulations.

All drug effects were designed to give the typical individual 10% decrease in HbA1c with 50 mg dose relative to placebo, implemented as maximum effect (E_{max}) functions. The half-maximal effective concentration was set to 0.03, matching the typical individuals' maximum concentrations after a dose of 25 mg. Corresponding E_{max} values were titrated to 0.455 (BASI), 0.8 (CLG), 0.465 (CLGI), 0.375 (EGP), and 1.18 (INCR) and were all associated with a 5% interindividual variability (IIV).

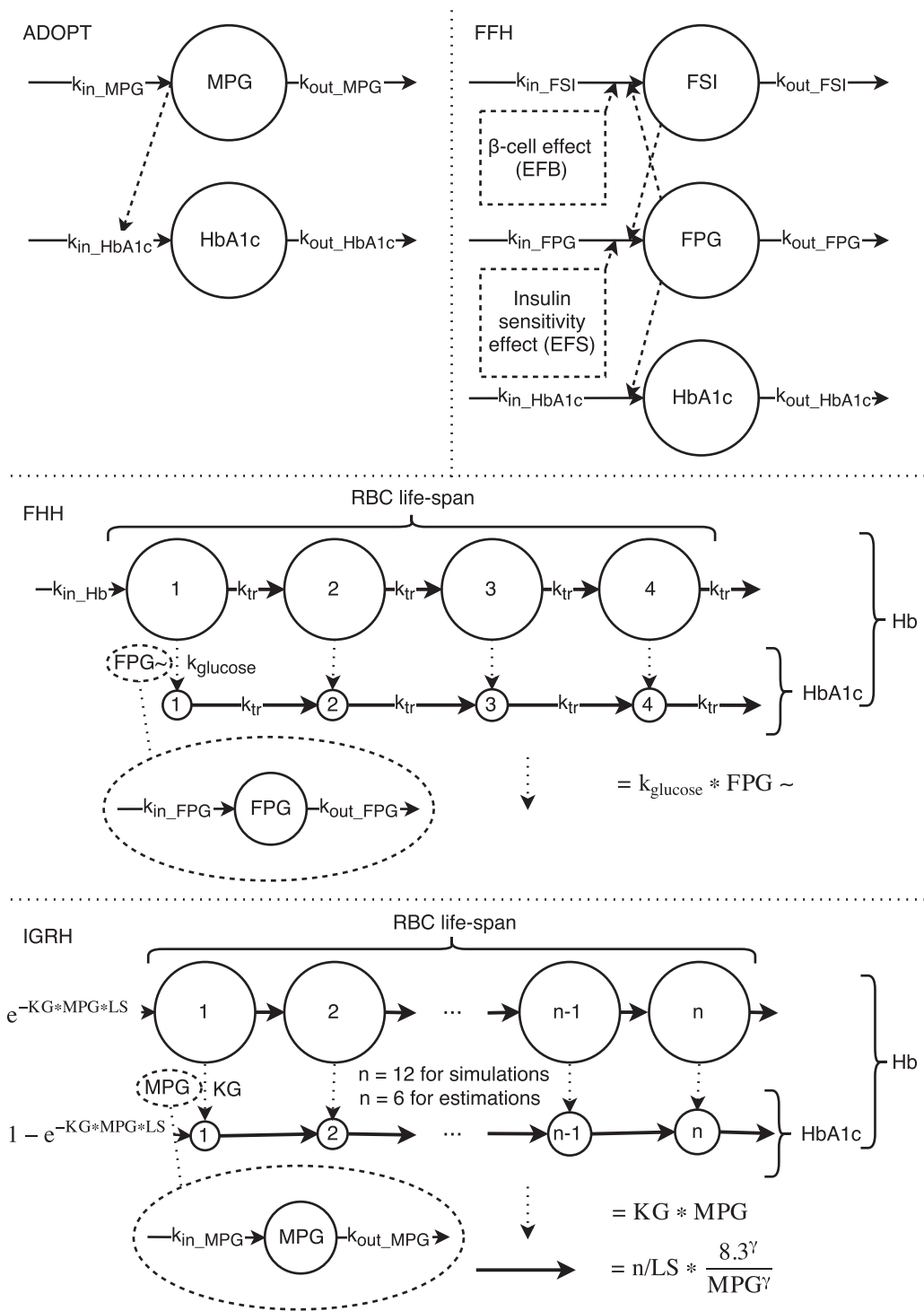


Figure 1 Schematic presentation of the A Dynamic HbA1c EndpOint Prediction Tool (ADOPT) model¹¹ (top left panel), the FPG-FSI-HbA1c model¹² (FFH; top right panel), the FPG-Hb-HbA1c model⁴ (FHH; middle panel) and the Integrated Glucose-RBC-HbA1c model¹³ (IGRH; bottom panel). FPG, fasting plasma glucose; FSI, fasting serum insulin; Hb, hemoglobin; HbA1c, glycosylated hemoglobin; KG, turnover of glucose; $k_{glucose}$, glycosylation rate constant; k_{in} , input rate constant; k_{out} , output rate constant; k_{tr} , transit rate constant; LS, life-span; MPG, mean plasma glucose; RBC, Red Blood Cells.

Output from the IGI model was coupled with the full IGRH model (24 compartments for RBCs) to produce HbA1c. An IIV of 17% was added on the proportional

HbA1c error. Simulation settings were chosen to render a sample size ≥ 10 patients per arm with a conventional t -test at 80% power with $\alpha = 0.05$.

Glucose, insulin, and HbA1c were sampled in the morning (after 11 hours of fasting) at weeks 0, 4, 8, and 12. The integrated 24-hour glucose exposure was used to calculate the MPG. Biomarker profiles are shown in **Supplementary Figures S3–S7**.

Estimations

The four HbA1c models were implemented as suggested by the authors of the original publications and a placebo model with IIV was added.

- In the ADOPT model, k_{in_HbA1c} and k_{out_HbA1c} were fixed to 0.0116 day^{-1} and 0.0323 day^{-1} from the published values, whereas all other parameters were estimated. Placebo affected k_{in_MPG} .
- In the FFH model, all parameters were estimated and a steady-state-solution of FPG and FSI was implemented for runtime and stability reasons. Placebo affected the turnover of glucose.
- In the FHH model, all parameters were estimated. Placebo affected k_{in_FPG} .
- In the IGRH model, all system parameters were fixed to their published values, whereas design parameters (k_{out_MPG} , MPG_0) were estimated. To reduce runtime, 12 transit compartments, instead of 24, were used and the system parameter turnover of glucose (KG) was re-estimated. The estimate was fixed to $1.42 \times 10^{-4} \text{ mM}^{-1} \times \text{day}^{-1}$ instead of the published $1.51 \times 10^{-4} \text{ mM}^{-1} \times \text{day}^{-1}$. Placebo affected k_{in_MPG} .

Drug model selection

Drug effects in the creation of data were E_{max} functions. However, that may not necessarily translate to E_{max} functions on FPG, MPG, or FSI. Therefore, the shape and position of the drug effect was investigated in each of the four estimation models. Linear, E_{max} , or sigmoidal E_{max} models were tested on all glucose/insulin parameters (i.e., k_{in} , k_{out} , KG, EFS, and/or EFB) and all combinations of these parameters. The likelihood ratio test was used to determine statistical significance using the difference in objective function value (ΔOFV) between the full (with drug effect) and the reduced model (without drug effect). If ΔOFV was the same for two competing drug effect models, a parsimonious approach was taken, benefiting smaller models. Drug effects on HbA1c were not investigated.

Power and sample size

The power to detect a drug effect was assessed as the ΔOFV between full and reduced models. Sample sizes were calculated with the Monte Carlo Mapped Power method (sampling from individual ΔOFV values) using the number of differing parameters between the full and reduced model as the degrees of freedom.²⁸ The sample size for a conventional *t*-test was also calculated at different levels of power (80% or 90%) and varying alpha (0.1%, 1%, 5%, or 10%), comparing the change from baseline at 12 weeks between placebo and 50 mg. The signal-to-noise ratio (SNR) was defined as the corresponding true effect size (difference in means, delta), back-calculated from the given alpha, power and sample sizes from the Monte Carlo Mapped Power, assuming a $SD = 1$. The SNR will be similar regardless of alpha or power, whereas sample sizes are heavily dependent on these values. The SNR for model-

based analysis corresponds to identifying a statistically significant concentration-effect relationship using all four arms (placebo and three doses), whereas the SNR for the *t*-test corresponds to identifying a statistically significant effect between placebo and the highest dose.

Prognosis (12 weeks)

The final parameter estimates of the full models were used to simulate 100 trials of 12 weeks duration with 600 patients, equally distributed among four dose arms. Simulations of baseline corrected HbA1c at 12 weeks of treatment were compared with the “observed” data, represented by baseline corrected HbA1c from the IGI-IGRH simulations at 12 weeks. Results were evaluated using the estimation error, which was calculated as the difference between prognosis and “observations” of the placebo-corrected change from baseline in HbA1c at 12 weeks.

Extrapolation (26 weeks)

The final parameter estimates were also used to make an extrapolation of the outcome of a 26-week (6-month) trial similarly as for 12-week prognoses. Simulations of baseline corrected HbA1c at 26 weeks of treatment were compared with “observations,” represented by baseline corrected HbA1c from the IGI-IGRH simulations at 26 weeks. Results were evaluated as estimation error at 26 weeks.

A schematic view of the study design is shown in **Supplementary Figure S2**.

Software

Data management, statistical calculations, and graphics were performed using R version 3.2.2 (R Core Team 2015, Vienna, Austria),²⁹ data simulation and estimation in NONMEM version 7.3 (Icon Development Solutions, Ellicott City, MD),³⁰ which was run through PsN version 4.5 (Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden).^{31,32}

RESULTS

Drug model selection

A summary of the results of the drug model selection are shown in **Supplementary Table S1**. Independent of the simulated site of action (SoA) of the drug in the IGI-IGRH model, the ADOPT, and IGRH models identified the same drug effect model: E_{max} function on k_{out_MPG} . In addition, the FHH model identified E_{max} functions in all cases. However, depending on the SoA, the drug effect was either identified on k_{in_FPG} or k_{out_FPG} .

For the FFH model, the most mechanistic model in terms of glucose and insulin, the best model varied greatly depending on the SoA. Thus, the FFH model could identify increased EFB when the drug effect had been placed on insulin secretion (BAS1 and INCR). For INCR, a drug effect was also identified on the turnover of glucose, likely related to the INCR effect mainly affecting PPG, which is not part of FPG and, thus, is difficult to model using FPG. When the original drug effect had been placed on CLGI, the FFH model identified an effect on insulin sensitivity plus an additional effect on EFB. This additional effect was a reduction of beta cell effect, reflecting the expected downregulation of

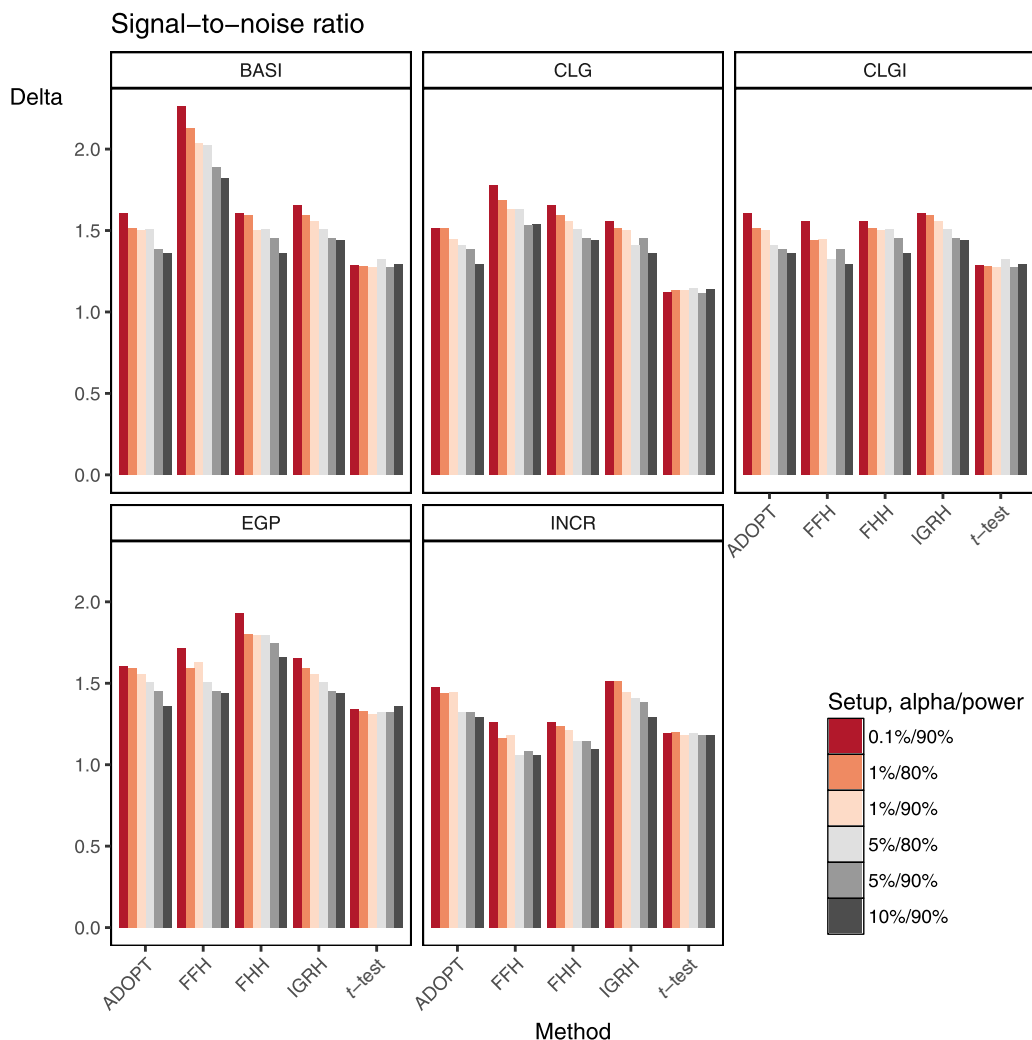


Figure 2 Signal-to-noise ratio (SNR) for site of drug action by model. Ratio is reported for several combinations of type 1 error rate (alpha) and power. For model-based analysis, the SNR corresponds to identifying a statistically significant concentration-effect relationship while for *t*-test, the SNR corresponds to identifying a statistically significant effect between placebo and the highest dose. ADOPT, A Dynamic HbA1c EndpOint Prediction Tool model; BASI, basal insulin secretion; CLG, insulin-independent glucose; CLGI, insulin-dependent glucose; EGP, endogenous glucose production; FFH, FPG-FSI-HbA1c model; FHH, FPG-Hb-HbA1c model; IGRH, Integrated Glucose-RBCCell-HbA1c model; INCR, incretin release.

insulin due to the smaller need for insulin, which is also mechanistically reasonable.

The FFH model describes the glucose-insulin homeostasis in the overnight fasting state and drug effects on CLG and EGP can either be effectuated via the KG or by improving insulin sensitivity (EFS) and/or insulin secretion, in which results indicate the latter explanation.

Power and sample size

In **Figure 2**, the SNR for the four HbA1c-models and the *t*-test are shown. The SNR is similar for the ADOPT and IGRH models, although consistently slightly lower for the ADOPT model compared to the IGRH model. The consistency is most likely due to the shared driver for HbA1c formation (i.e., MPG). Comparing the two models using FPG to drive HbA1c formation, results are less consistent as the FFH model also makes use of insulin. Insulin observations

strengthen the SNR when the drug effect is related to insulin secretion; however, the restriction imposed on the model-relating insulin and glucose observations through the HOMA equations rather seems to dilute the SNR when the drug effect is not directly related to insulin. The FFH model has the highest SNR of all models when the drug effect is stimulating BASI, whereas the FHH model has the highest SNR of all models when the drug inhibits EGP. The *t*-test generally has the lowest SNR, although for drugs stimulating INCR, it is on par with the worse models.

Prognosis (12 weeks) and extrapolation (26 weeks)

There was, in general, a good agreement between observed and simulated HbA1c change from baseline for all models, as shown in **Figure 3** and **Figure 4**. Independent of where the drug effect had its SoA, at what time the effect was assessed, and which model was used,

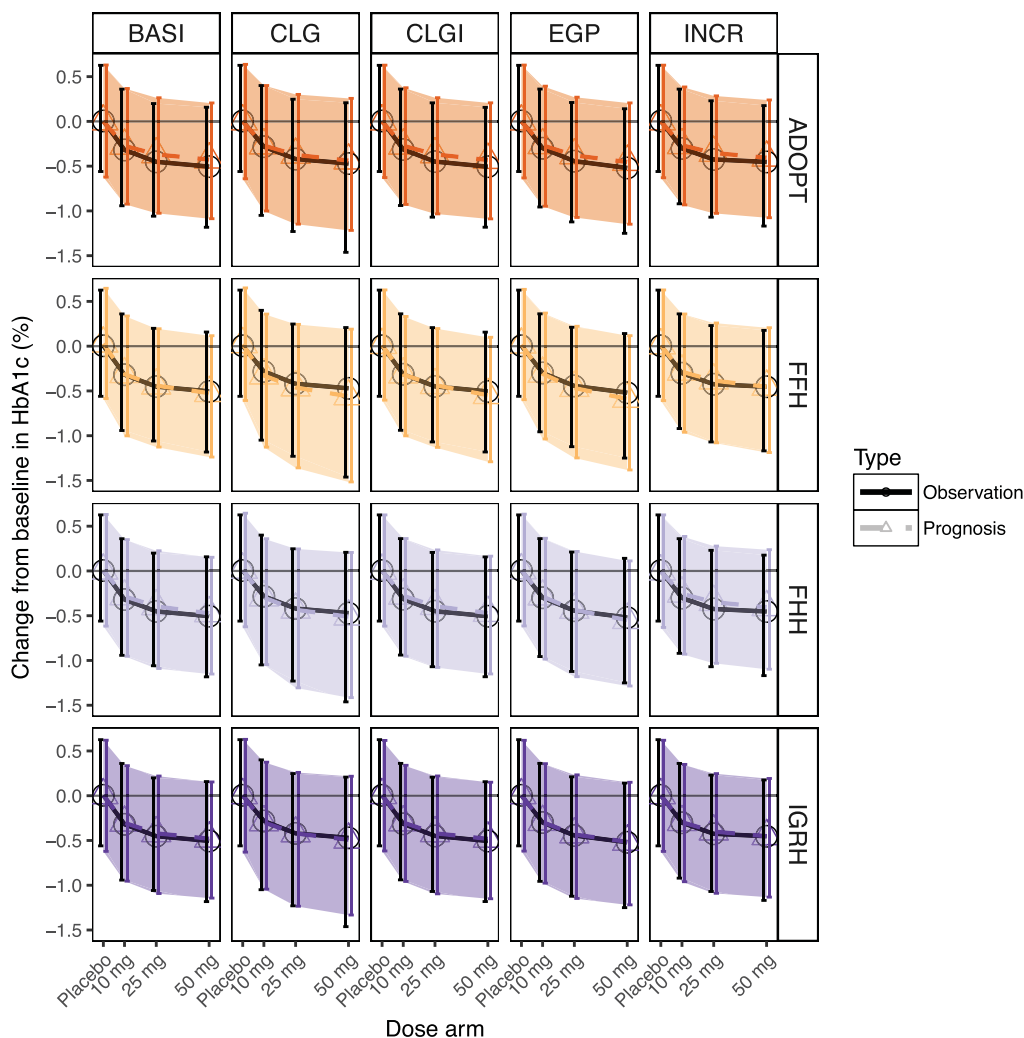


Figure 3 Baseline-corrected change in HbA1c at 12 weeks by dose. The mean and the corresponding 95% confidence interval is shown for observations (open circles, black solid line, and bars) and model simulations based on data up to 12 weeks (open triangles, shaded area, colored dashed lines and bars). ADOPT, A Dynamic HbA1c EndpOint Prediction Tool model; BASI, basal insulin secretion; CLG, insulin-independent glucose; CLGI, insulin-dependent glucose; EGP, endogenous glucose production; FFH, FPG-FSI-HbA1c model; FHH, FPG-Hb-HbA1c model; IGRH, Integrated Glucose-RBC-HbA1c model; INCR, incretin release.

differences were small both on main trend and confidence intervals.

Figure 5, depicting the differences between prognoses and observations of placebo-corrected and baseline-corrected HbA1c, shows that all models had good 12-week prognostic properties, with the IGRH being the most accurate across all SoAs.

The accuracy was even better for the 26-week extrapolations (**Figure 6**), with both the IGRH and ADOPT model extrapolating well. The accuracy of the extrapolations from the FHH model was improved compared to 12 weeks for all drug SoAs, except CLG and EGP, which were slightly biased. The FFH performed poorly in the extrapolations with higher bias across all SoAs compared to prognoses, with a severe bias for the extrapolations for CLG and EGP.

A summary of how the models performed in all the investigated aspects is shown in **Table 1**.

DISCUSSION

Drug model selection

The E_{max} models were used when simulating data using the IGI model. The nonlinearity of the drug effect was conserved also when glucose was modeled using less complex systems. The ADOPT and IGRH models were consistent across drug effects and identified E_{max} models on glucose elimination (k_{out}) throughout. With FFH, combinations of effects were needed to explain the full drug effect. However, all models identified at least one relationship to be nonlinear. The FHH model consistently identified E_{max} models, however, on different system parameters.

The FFH is the only of the tested models with a semimechanistic model for glucose and insulin. Although the glucose model is simplified compared to the IGI model, it could correctly identify the various sites of drug effects, based on statistical criteria (not physiological rationale). For example,

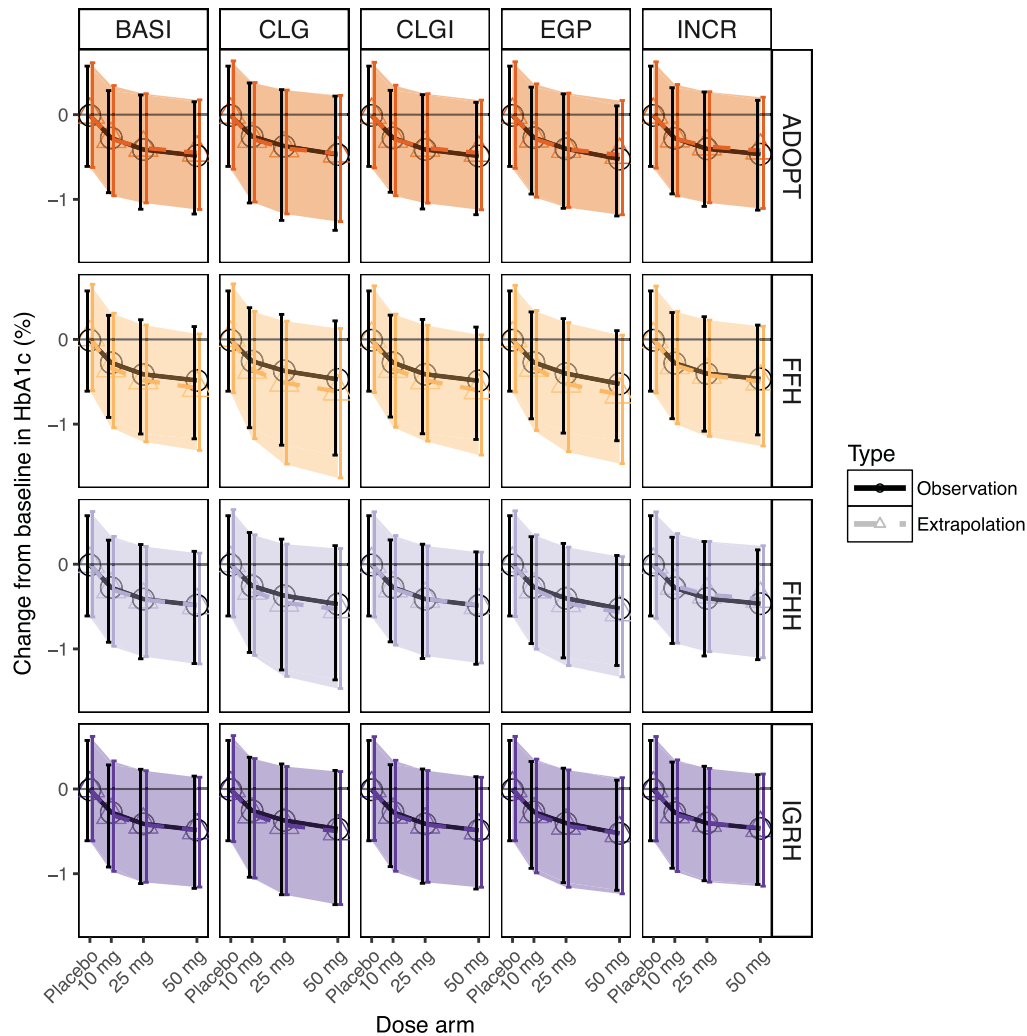


Figure 4 Baseline-corrected changes in HbA1c at 26 weeks by dose. The mean and the corresponding 95% confidence interval is shown for observations (open circles, black solid line, and bars) and model simulations based on data up to 12 weeks (open triangles, shaded area, colored dashed lines, and bars). ADOPT, A Dynamic HbA1c EndPoint Prediction Tool model; BASI, basal insulin secretion; CLG, insulin-independent glucose; CLGI, insulin-dependent glucose; EGP, endogenous glucose production; FFH, FPG-FSI-HbA1c model; FHH, FPG-Hb-HbA1c model; IGRH, Integrated Glucose-RBC-HbA1c model; INCR, incretin release.

when the drug effect in the IGI model was on BASI, the identified drug effect was on EFB and a drug effect in the IGI model on CLG was identified as a drug effect on EFS in the FFH model. Thus, this model offers an additional benefit over the others in identifying the SoA of the drug.

Power and sample size

The MPG-driven models (ADOPT and IGRH) performed best with drug-affecting INCR. This is expected as INCR affects PPG and the information about PPG is contained in MPG, not FPG. The FFH model has the highest power to detect drug effects acting on BASI. This suggests that sampling insulin is useful for insulin secretagogues and the FFH model is more highly powered than the FHH model, because it has additional information given from FSI. The difference between FPG and MPG as a driver for HbA1c is relatively small, suggesting that, in many cases, a fasting

value can be used instead of an integrated 24-hour mean value without loss of power, especially for drug effects on EGP. The SNR for any model-based approach is generally higher than for a regular *t*-test. Again, this is expected because the *t*-test disregards information from intermediate dose arms, midpoint samples, and glucose; focusing only on 12 weeks of HbA1c for placebo and highest dose arm. The FFH and FHH models perform worse than the *t*-test for a drug stimulating INCR, likely because the models cannot fit the flat FSI and FPG profiles simultaneously with lowered HbA1c with drug effects only on glucose/insulin parameters (see **Supplementary Figure S3**). Adding drug effects on HbA1c parameters resolved this issue (results not shown). In addition, the FFH model is penalized due to the number of parameters. The correct model choice can reduce the size of a clinical trial, because power determines the sample size required.

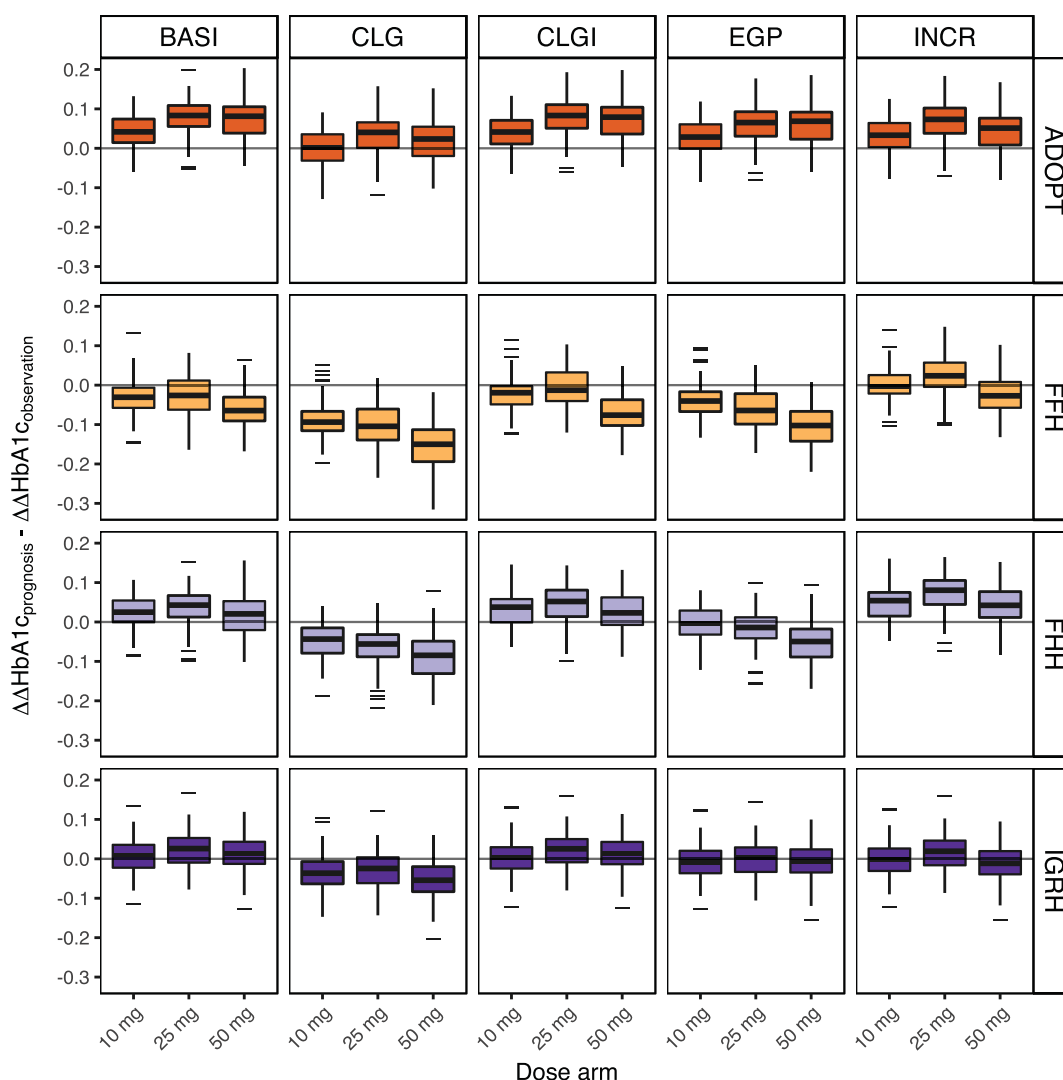


Figure 5 Placebo-corrected, baseline-corrected HbA1c at 12 weeks by dose. ADOPT, A Dynamic HbA1c EndPoint Prediction Tool model; BASI, basal insulin secretion; CLG, insulin-independent glucose; CLGI, insulin-dependent glucose; EGP, endogenous glucose production; FFH, FPG-FSI-HbA1c model; FHH, FPG-Hb-HbA1c model; IGRH, Integrated Glucose-RBC-HbA1c model; INCR, incretin release.

Prognosis

All models make good prognoses, with small deviations from observed values. Because we are fitting the models to 12-week data this is unsurprising. For prognostic purposes, the model choice will not affect results greatly. Thus, a less mechanistic model, such as ADOPT or FFH, may be used if it has other advantages.

Extrapolation

The extrapolation results largely follow similar trends as the prognoses. The main difference from the 12-week prognoses is that the FFH model is slightly overpredicting the drug effect. This is likely due to an incorrect shape of the model fit, in which the time to steady-state is too long. Hence, the FFH model should not be the first choice when making extrapolations to 26 weeks.

Additional considerations

For convenience, we assumed all patients received standardized meals with regular intervals and followed study protocol. These assumptions give a less variable population than normally expected from a 12-week HbA1c study. To account for this, we allowed an inflated IIV in both the pharmacokinetic and pharmacodynamic parameters plus noise in terms of IIV on residual error for HbA1c. The HbA1c residual error was inflated so at least 10 individuals per arm in a conventional *t*-test would be required for 80% power. Normally, a phase IIb trial requires even more patients for 80% power, and part of this high power with few individuals is related to the study assumptions (standardized meals, no dropouts). However, we focused on finding the properties that distinguish the proposed models from one another and also from conventional practice in

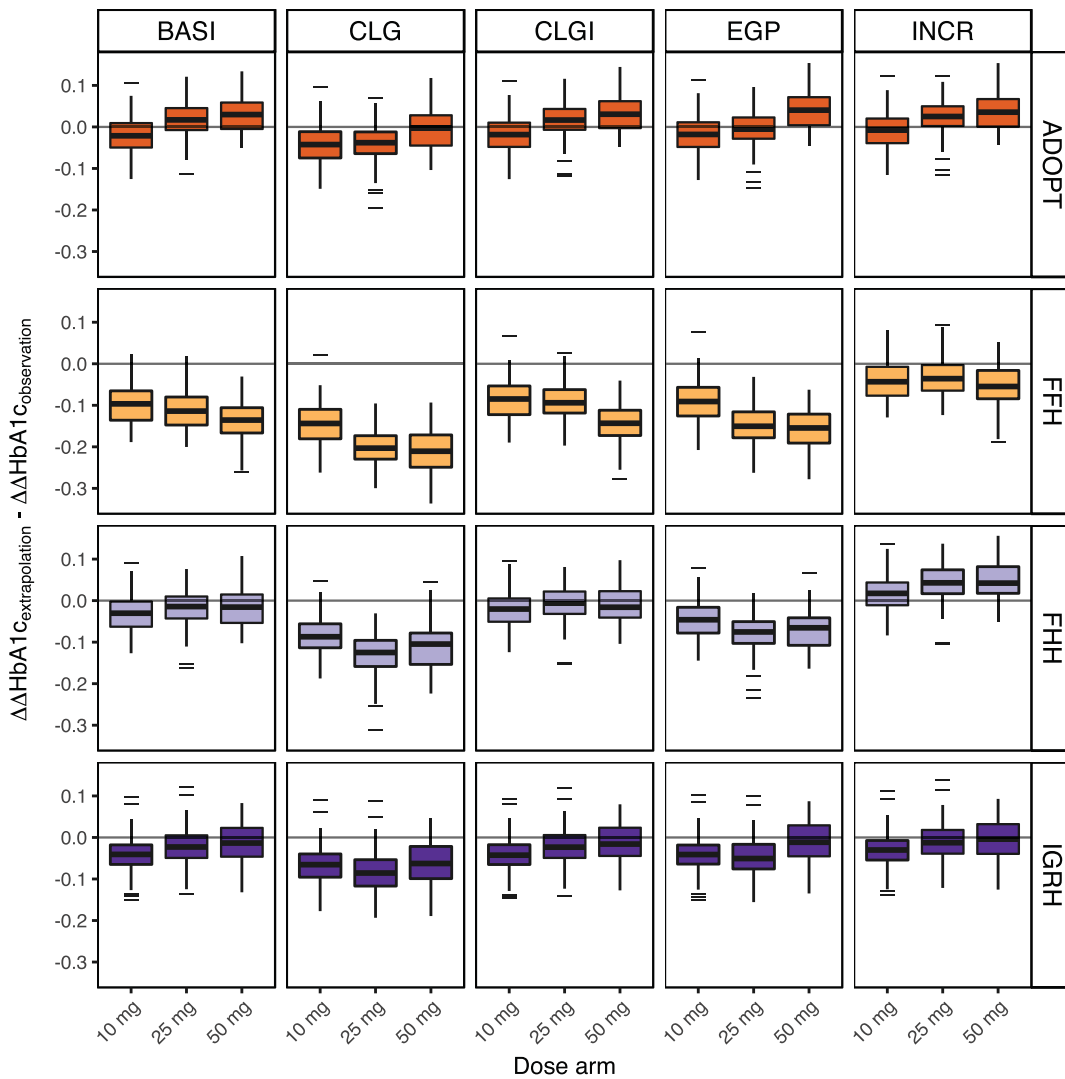


Figure 6 Placebo-corrected, baseline-corrected HbA1c at 26 weeks by dose. ADOPT, A Dynamic HbA1c EndPoint Prediction Tool model; BASI, basal insulin secretion; CLG, insulin-independent glucose; CLGI, insulin-dependent glucose; EGP, endogenous glucose production; FFH, FPG-FSI-HbA1c model; FHH, FPG-Hb-HbA1c model; IGRH, Integrated Glucose-RBC-HbA1c model; INCR, incretin release.

quantitative assessment of clinical data. Thus, the errors added should distort the signal enough to make the conclusions valid. This is supported by the fact that the SNR is consistent across different levels of alpha and power for each model.

There are both strengths and weaknesses associated with using simulations for the comparison between the models. The advantage of simulating the data is the control of the true relationships between drug effect and the glucose system, but the validity of the results will depend on how accurately the simulation model represents the true nature of the biological system. The IGI model has been shown to perform well in analyzing glucose challenges with and without drug effect.^{16,33,34} It was, however, developed for a short-term glucose test and not 12-week studies of glucose and may lack certain aspects of the glucose homeostasis needed to accurately describe glucose over

an extended time period. Kjellsson *et al.*⁸ showed that the IGI model performed reasonably well in predicting HbA1c over a 12-week period for drug effect on EGP and BASI, however, how well the model performs for other mechanisms of actions has not been shown and is speculative.

As recommended in the original paper, k_{out_HbA1c} in the ADOPT model was fixed to the reported value.¹¹ This allows less flexibility than the IGRH model, which could explain the slightly lower power and slightly higher inaccuracy in 12-week prognosis observed compared to the IGRH model. The predictive performance of the IGRH and ADOPT models has previously been investigated head-to-head.⁹ In that study, the IGRH model was used with fixed parameters, results slightly favoring the ADOPT model. This shows that these models are similar in performance.

The FFH model was implemented with a steady-state solution of FPG and FSI to decrease runtimes and for

Table 1 Results summary

Model	SoA identification	Power	Prognosis (%)	Extrapolation (%)	Relative runtime
ADOPT	No		1.12	1.10	1.4
FFH (steady state)	Yes	BASI, CLG	1.20	1.54	1 ^a
FFH (differential equations)	–	–	–	–	2.6 ^b
FHH	No	EGP	1.30	1.36	6.8
IGRH (12 transit compartments)	No	CLGI, INCR	1.08	1.19	150
IGRH (24 transit compartments)	–	–	–	–	2.0 ^b

The table summarizes the four aspects compared for the four models (ADOPT, FFH, FHH, and IGRH): SoA identification (dichotomous: yes/no), power (SoA with the highest power), prognosis (root mean-squared error (RMSE) of $\Delta\Delta\text{HbA1c}$) and extrapolation (RMSE of $\Delta\Delta\text{HbA1c}$). In addition, relative runtimes (reference fastest model: FFH) is included, listing also the full models: FFH with differential equations and IGRH with 24 transit compartments.

ADOPT, A Dynamic HbA1c EndPoint Prediction Tool model; BASI, basal insulin secretion; CLG, insulin-independent glucose; CLGI, insulin-dependent glucose; EGP, endogenous glucose production; FFH, FPG-FSI-HbA1c model; FHH, FPG-Hb-HbA1c model; IGRH, Integrated Glucose-RBC-HbA1c model; INCR, incretin release; SoA, site of action.

^a30 minutes: drug effect – BASI; cluster specification: OS – Linux 6.3, compiler – gfortran 5.1.1 from GCC, processor – Intel Xeon E5645.

^bRelative to simplified model.

stability reasons (could not reach successful minimization). This will only marginally affect our investigations as the changes in glucose were instantaneous and effect steady-state on glucose and insulin would have been reached in between the first and second observations. However, it should be noted that if the effect of steady-state on glucose and insulin is not reached between the first and the second observations, due to a delayed concentration-effect relationship, the full FFH should be used.

The study design was biased toward the IGRH model, because this model was used to simulate HbA1c data. The IGRH model represents the most mechanistic description of current knowledge of physiology and glycation of hemoglobin and, hence, the most appropriate choice for simulating HbA1c. Unsurprisingly, this model showed the most consistent estimates of HbA1c. However, a model misspecification was intentionally introduced between creation and analysis when the number of RBC compartments was reduced, and, thus, the impact of the favor toward IGRH was decreased.

General

This work does not include all but rather a representative selection of published HbA1c models. The simplest HbA1c model, in which linked indirect response models describe both FPG and HbA1c, is not explicitly presented. However, in the FHH model, FPG is described with a regular turnover model and, in the FFH model, FPG is linked to a regular indirect response model of HbA1c, although the FPG model is more mechanistic. Thus, the properties of this simplest model can be derived from the FHH and FFH models. In identifying drug effects, this simplest model should behave as the FHH model with a good ability to identify the nonlinearity of glucose with effects on k_{in} or k_{out} depending on drug effect SoA. In terms of power, the behavior should resemble that of the FHH for the drug effect related to insulin and that of FFH for drug effects unrelated to insulin. This model is, although better than a *t*-test, not in the top of the investigated models. As prognosis was good for all models, there is no reason to believe otherwise for the simplest model, however, it is likely to incur the same extrapolation problems as the FFH model: underpredicting the time

to steady-state from 12-week data, consequently overpredicting the drug effect at week 26.

Final remarks

One extension of this work would be to investigate combinations of drug effects, as many drugs have several mechanisms of action. For example, metformin, the current standard-of-care, has been hypothesized to affect both EGP and CLGI. The power to distinguish both primary and secondary drug effects also applies to add-on therapies.

All models investigated in this work performed well in one or several aspects (**Table 1**). For many purposes, the model choice will not impact results greatly but for certain drug SoAs there is lots to be gained in terms of identification, power, and predictive performance. The model choice will also guide what biomarkers to sample.

A good source for published models is the Drug Disease Model Resource (DDMoRe) model repository (<http://repository.ddmore.eu>).

Acknowledgments. The authors thank Bengt Hamrén and Magnus Åstrand at AstraZeneca, Mölndal, for valuable input.

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 the European Federation of Pharmaceutical Industries and Associations (EFPIA) companies' in kind contribution. The DDMoRe project is also financially supported by contributions from Academic and SME partners. Financial support was also received from the IDeAI Project (FP7 HEALTH 2013-602552).

Conflict of Interest. The authors declared no conflict of interest.

Author Contributions. G.J.W., M.O.K., and M.C.K. wrote the manuscript. G.J.W., M.O.K., and M.C.K. designed the research. G.J.W., M.O.K., and M.C.K. performed the research. G.J.W., M.O.K., and M.C.K. analyzed the data.

1. Ajmera, I., Swat, M., Laibe, C., Le Novère, N. & Chelliah, V. The impact of mathematical modeling on the understanding of diabetes and related complications. *CPT Pharmacometrics Syst. Pharmacol.* **10**, e54 (2013).

2. Zhang, X., Schneck, K., Bue-Valleskey, J., Yeo, K.P., Heathman, M. & Sinha, V. Dose selection using a semi-mechanistic integrated glucose-insulin-glucagon model: designing phase 2 trials for a novel oral glucokinase activator. *J. Pharmacokinet. Pharmacodyn.* **40**, 53–65 (2013).
3. Luu, K.T. *et al.* Population pharmacokinetics and pharmacodynamics of IONIS-GCGRX, an antisense oligonucleotide for type 2 diabetes mellitus: a red blood cell lifespan model. *J. Pharmacokinet. Pharmacodyn.* **44**, 179–191 (2017).
4. Hamrén, B., Björk, E., Sunzel, M. & Karlsson, M.O. Models for plasma glucose, HbA1c, and hemoglobin interrelationships in patients with type 2 diabetes following tesaglitazar treatment. *Clin. Pharmacol. Ther.* **84**, 228–235 (2008).
5. Claussen, A. *et al.* Impact of demographics and disease progression on the relationship between glucose and HbA1c. *Eur. J. Pharm. Sci.* **104**, 417–423 (2017).
6. Stringer, F., DeJongh, J., Scott, G. & Danhof, M. A model-based approach to analyze the influence of UGT2B15 polymorphism driven pharmacokinetic differences on the pharmacodynamic response of the PPAR agonist sipoglitazar. *J. Clin. Pharmacol.* **54**, 453–461 (2014).
7. de Winter, W. *et al.* Dynamic population pharmacokinetic-pharmacodynamic modelling and simulation supports similar efficacy in glycosylated haemoglobin response with once or twice-daily dosing of canagliflozin. *Br. J. Clin. Pharmacol.* **83**, 1072–1081 (2017).
8. 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).
9. Møller, J.B., Kristensen, N.R., Klim, S., Karlsson, M.O., Ingwersen, S.H. & Kjellsson, M.C. Methods for predicting diabetes phase III efficacy outcome from early data: superior performance obtained using longitudinal approaches. *CPT Pharmacometrics Syst. Pharmacol.* **3**, e122 (2014).
10. Liu, D. *et al.* Translational modeling and simulation in supporting early-phase clinical development of new drug: a learn-research-confirm process. *Clin. Pharmacokinet.* **56**, 925–939 (2017).
11. Møller, J.B. *et al.* Longitudinal modeling of the relationship between mean plasma glucose and HbA1c following antidiabetic treatments. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e82 (2013).
12. de Winter, W. *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.* **33**, 313–343 (2006).
13. Lledó-García, R., Mazer, N.A. & Karlsson, M.O. A semi-mechanistic model of the relationship between average glucose and HbA1c in healthy and diabetic subjects. *J. Pharmacokinet. Pharmacodyn.* **40**, 129–142 (2013).
14. Matthews, D.R., Hosker, J.P., Rudenski, A.S., Naylor, B.A., Treacher, D.F. & Turner, R.C. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* **28**, 412–419 (1985).
15. Lledó-García, R., Kalicki, R.M., Uehlinger, D.E. & Karlsson, M.O. Modeling of red blood cell life-spans in hematologically normal populations. *J. Pharmacokinet. Pharmacodyn.* **39**, 453–462 (2012).
16. 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).
17. Ginsberg, B.H. Factors affecting blood glucose monitoring: sources of errors in measurement. *J. Diabetes Sci. Technol.* **3**, 903–913 (2009).
18. Marcovina, S. *et al.* Standardization of insulin immunoassays: report of the American Diabetes Association Workgroup. *Clin. Chem.* **53**, 711–716 (2007).
19. Amin, N.B., Wang, X., Jain, S.M., Lee, D.S., Nucci, G. & Rusnak, J.M. Dose-ranging efficacy and safety study of ertugliflozin, a sodium-glucose co-transporter 2 inhibitor, in patients with type 2 diabetes on a background of metformin. *Diabetes Obes. Metab.* **17**, 591–598 (2015).
20. Ross, S.A., Rafeiro, E., Meinicke, T., Toorawa, R., Weber-Born, S. & Woerle, H.-J. Efficacy and safety of linagliptin 2.5 mg twice daily versus 5 mg once daily in patients with type 2 diabetes inadequately controlled on metformin: a randomised, double-blind, placebo-controlled trial. *Curr. Med. Res. Opin.* **28**, 1465–1474 (2012).
21. Wilding, J.P.H., Leonsson-Zachrisson, M., Wessman, C. & Johnsson, E. Dose-ranging study with the glucokinase activator AZD1656 in patients with type 2 diabetes mellitus on metformin. *Diabetes Obes. Metab.* **15**, 750–759 (2013).
22. Ikeda, S. *et al.* A novel and selective sodium-glucose cotransporter-2 inhibitor, tofogliflozin, improves glycaemic control and lowers body weight in patients with type 2 diabetes mellitus. *Diabetes Obes. Metab.* **17**, 984–993 (2015).
23. Gastaldelli, A., Nauck, M.A. & Balena, R. Eight weeks of treatment with long-acting GLP-1 analog taspoglutide improves postprandial insulin secretion and sensitivity in metformin-treated patients with type 2 diabetes. *Metabolism* **62**, 1330–1339 (2013).
24. Riggs, M.M. *et al.* Population pharmacokinetics of empagliflozin, a sodium glucose cotransporter 2 inhibitor, in patients with type 2 diabetes. *J. Clin. Pharmacol.* **53**, 1028–1038 (2013).
25. Sloan-Lancaster, J. *et al.* Double-blind, randomized study evaluating the glycemic and anti-inflammatory effects of subcutaneous LY2189102, a neutralizing IL-1 antibody, in patients with type 2 diabetes. *Diabetes Care* **36**, 2239–2246 (2013).
26. Kovacs, C.S. *et al.* Empagliflozin improves glycaemic and weight control as add-on therapy to pioglitazone or pioglitazone plus metformin in patients with type 2 diabetes: a 24-week, randomized, placebo-controlled trial. *Diabetes Obes. Metab.* **16**, 147–158 (2014).
27. Ferrannini, E., Seman, L., Seewaldt-Becker, E., Hantel, S., Pinnetti, S. & Woerle, H.J. A phase IIb, randomized, placebo-controlled study of the SGLT2 inhibitor empagliflozin in patients with type 2 diabetes. *Diabetes Obes. Metab.* **15**, 721–728 (2013).
28. 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).
29. R Core Team. R: a language and environment for statistical computing. [Internet]. R Foundation for Statistical Computing, Vienna, Austria. <<http://www.R-project.org/>> (2014).
30. Beal, S., Sheiner, L.B., Boeckmann, A. & Bauer, R.J. *NONMEM User's Guides. (1989–2009)* (Icon Development Solutions, Ellicott City, MD, 2009).
31. Lindbom, L., Ribbing, J. & Jonsson, E.N. Perl-speaks-NONMEM (PsN)—a Perl module for NONMEM related programming. *Comput. Methods Programs Biomed.* **75**, 85–94 (2004).
32. Keizer, R.J., Karlsson, M.O. & Hooker, A. Modeling and simulation workbench for NONMEM: tutorial on Pirana, PsN, and Xpose. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e50 (2013).
33. 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).
34. Schneck, K.B., Zhang, X., Bauer, R., Karlsson, M.O. & Sinha, V.P. Assessment of glycemic response to an oral glucokinase activator in a proof of concept study: application of a semi-mechanistic, integrated glucose-insulin-glucagon model. *J. Pharmacokinet. Pharmacodyn.* **40**, 67–80 (2013).

© 2018 The Authors CPT: Pharmacometrics & Systems Pharmacology published by Wiley Periodicals, Inc. on behalf of American Society for Clinical Pharmacology and Therapeutics. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

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